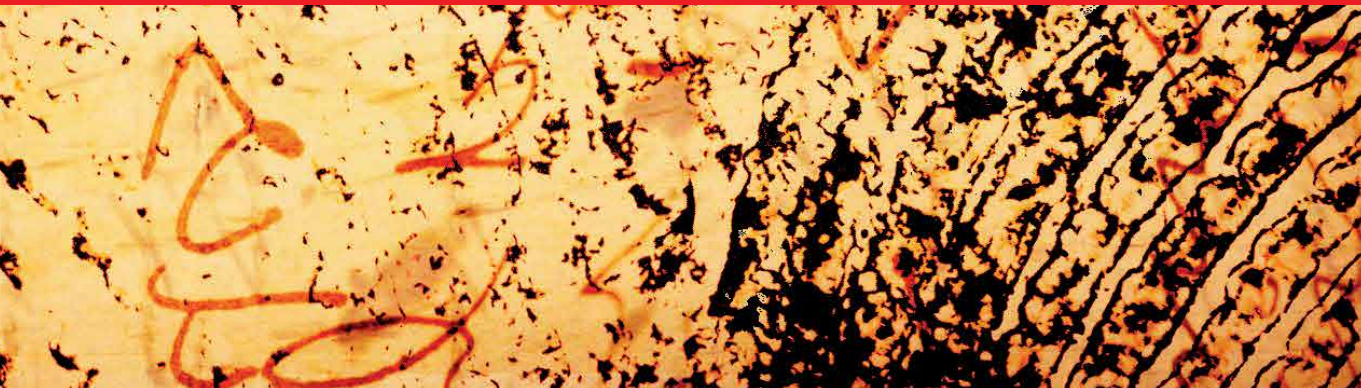
IntechOpen

Digital Forensic Science

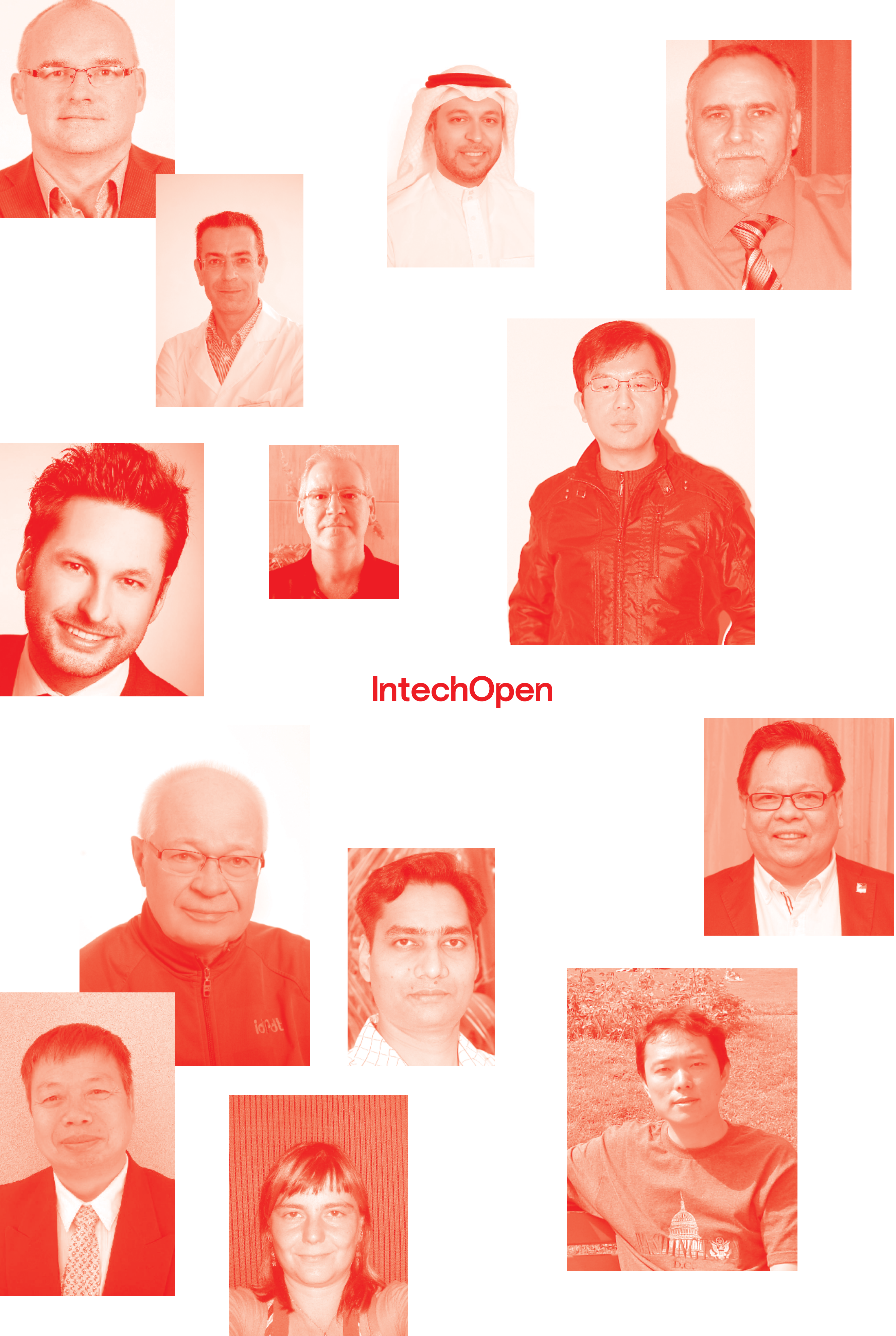
Edited by B Suresh Kumar Shetty
and Pavanchand Shetty H



Digital Forensic Science

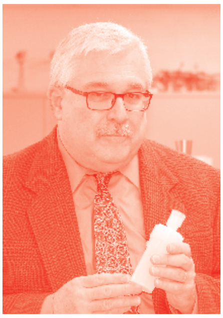
Edited by B Suresh Kumar Shetty and Pavanchand Shetty H

Published in London, United Kingdom



IntechOpen

Supporting open minds since 2005



Digital Forensic Science
http://dx.doi.org/10.5772/intechopen.78450
Edited by B Suresh Kumar Shetty and Pavanchand Shetty H

Contributors
Rajasree Thanka Raja, Mary Saira Bhanu S, Vladimir Ivanovich Vasilyev, Alexey Vulfin, Liliya Chernyakhovskaya, Salman Iqbal, Soltan Alharbi, Thorsten Floren M.A., Louise Kelly, Swati Sachan, Lei Ni, Fatima Almaghrabi, Richard Allmendinger, Yu-Wang Chen, Petra Perner, Rinaldi Munir, Harlili Harlili, Deepa Salian, Sofia Khatun

Notice
Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2020 by IntechOpen
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales,
registration number: 11086078, 5 Princes Gate Court, London, SW7 2QJ, United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Digital Forensic Science
Edited by B Suresh Kumar Shetty and Pavanchand Shetty H
p. cm.
Print ISBN 978-1-83880-259-2
Online ISBN 978-1-83880-260-8
eBook (PDF) ISBN 978-1-83968-742-6

Meet the editors

Dr. B. Suresh Kumar Shetty is a Professor in the Department of Forensic Medicine at Kasturba Medical College, Mangalore, MAHE, Manipal. He received his Master's degree at KMC, Manipal, and joined as faculty at KMC, Mangalore, where he has been teaching since 2005. He is presently appointed as the Honorary State Medico-Legal Consultant for three districts [Dakshina Kannada, Udupi and Coorg], Government of Karnataka. He has received his certificate in "Analytical Toxicology & Forensic DNA Typing" awarded by the Department of Analytical Toxicology, Amrita Institute of Medical Sciences, Cochin in November 2006, as well as his PG certificate in Torture Medicine [IMA AKN Sinha Institute, Patna] in 2011. He completed his MBA [Hospital Administration] from Sikkim Manipal University in 2016. He completed his FAIMER fellowship from MAHE, Manipal in 2018. He has co-authored chapters in 3 books and peer-reviewed papers for 15 reputed national and international journals, contributed articles in local and national newspapers, guided postgraduates in Forensic Medicine and undergraduates in ICMR student projects. He has published more than 80 papers, all national and international research papers in scientific journals. He has organized a number of continued medical education [CME] programs, police training programs, seminars and conferences in the capacity of Organizing Secretary from 2009-2019. He hosted the IAFM National conference in 2013. He was the Organizing Secretary of the 1st Indo-French Forensic International Congress in 2018 and also successfully conducted the 2nd Indo-French Congress in Lyon France as Co-Organizing Secretary in 2019. He has successfully completed international collaborations with Mekelle University and Lyon University, France. He is the advisor for the Centre of Forensic Odontology and Head of the Students wing of Bio-Ethics, KMC, Mangalore. He has attended scientific sessions in India as well as abroad like Malaysia, Singapore, Vietnam, Thailand, United States of America, Hong Kong, Australia, China, Russia, and France. He is the author of "Atlas Book on Forensic Pathology" published by Jayeepe Publisher. Forensic Analysis – Death to Justice by IntechOpen publisher, as well as chapters in books from Nova publisher and IntechOpen publisher. He was nominated by the International Bibliographical Centre, Cambridge, England, Selection Committee and earned a position of TOP 100 Health Professionals in 2009, who have made a significant contribution in their field to engender influence on a local, national and international issue and Individual Member of Sydney Forensic Medicine & Science Network, Australia and Asia Pacific Association of Medical Toxicology. He is an active member and Chairman of a registered Non-Governmental Organization "BELAKU" with a desire and wish to spread light among the youths to change society in the right direction.

Dr. Pavanchand Shetty H is an Associate Professor in the Department of Forensic Medicine, Kasturba Medical College, Mangalore, MAHE, Manipal. He received his Master's Degree at KMC, Manipal and joined as faculty at KMC, Mangalore where he has been teaching since 2011. He is presently appointed as Honorary District Medico-Legal consultant for Dakshina Kannada, Government of Karnataka. He has co-authored a chapter in a book titled NACP-FMT Practical Medico Legal Manual. He has guided postgraduates in Forensic Medicine and undergraduates in research projects. He has published more than 25 papers, as well as national and international research papers in scientific journals. He has organized a number of continued medical education (CME) programs, police training programs, seminars and conferences as a part of an organizing team. He was part of the organizing team of the 1st Indo-French Forensic International Congress in 2018 and also successfully conducted the 2nd Indo-French congress in Lyon France as part of organizing team in 2019. He is part of the training members of the Centre of Forensic Odontology. He has attended scientific sessions in India as well as abroad.

Contents

Preface

It is our pleasure to place before you the book *Digital Forensic Science*. This book makes up a major part of the broad specialty of Digital Forensic Science, comprising mainly of tools and technologies of cyber forensic experts for their future practice. This book is designed to merge a range of new ideas and unique works of authors from topics like fundamental principles of forensic cyber analysis, and protocols and laws related to the digital world. This information is very much-needed for the best of digital forensics. We hope that it will be useful to practitioners of forensic medicine, experts, cyber experts, law makers, investigating authorities, and undergraduate and postgraduate medical school graduates of medicine.

The experienced and enthusiastic authors have presented many ideas and innovative approaches. They have given a new outlook to this book and most importantly enabled us to grow and push through many hurdles, in our minds and in the process, and also learn how to edit and publish. We are truly proud of this book.

We wish to express our solemn sentiments and sincere thanks to Dr. Aditi S Shetty, Associate Professor, Department of Obstetrics and Gynecology from Kasturba Medical College, Mangalore, Manipal Academy of Higher Education (M.A.H.E), Manipal for her valuable feedback and suggestions while editing this book.

We wholeheartedly thank Mr. Mateo Pulko, Author Service Manager, and Ms. Sandra Bakic, Senior Commissioning Editor of IntechOpen, for their constant support and suggestions during the process of editing and thank the entire team of IntechOpen publisher for giving us the unique opportunity in this process. We would like to place our gratitude to our university, the Manipal Academy of Higher Education, Manipal.

B. Suresh Kumar Shetty and Pavanchand Shetty H.
Manipal Academy of Higher Education,
India

Section 1

Digital Forensics - Computer and Network

Chapter 1

Advancing Automation in Digital Forensic Investigations Using Machine Learning Forensics

Salman Iqbal and Soltan Abed Alharbi

Abstract

In the last few years, most of the data such as books, videos, pictures, medical and even the genetic information of humans are moving toward digital formats. Laptops, tablets, smartphones and wearable devices are the major source of this digital data transformation and are becoming the core part of our daily life. As a result of this transformation, we are becoming the soft target of various types of cybercrimes. Digital forensic investigation provides the way to recover lost or purposefully deleted or hidden files from a suspect's device. However, current man power and government resources are not enough to investigate the cybercrimes. Unfortunately, existing digital investigation procedures and practices require huge interaction with humans; as a result it slows down the process with the pace digital crimes are committed. Machine learning (ML) is the branch of science that has governs from the field of AI. This advance technology uses the explicit programming to depict the human-like behaviour. Machine learning combined with automation in digital investigation process at different stages of investigation has significant potential to aid digital investigators. This chapter aims at providing the research in machine learning-based digital forensic investigation, identifies the gaps, addresses the challenges and open issues in this field.

Keywords: digital forensic investigation, machine learning, evidence extraction, cybercrimes, automated data extraction

1. Introduction

Worldwide usage of mobile smart devices has increased dramatically over the past two decades and is becoming the part of our daily life. The term smart device ranges from variety of devices that includes mobile phones, smartphones, tablets, GPS and so on. The popularity of these smart devices is increased significantly due to their processing power, huge storage capabilities and less cost. Consequently, they can hold the enormous amount of commercial and private user's data. These devices are the essential part of our daily life because they contain private and essential information of users. However, these devices are also vulnerable to attackers and are often becoming the major part of criminal's activities, IP theft, intrusions, security threats, accidents reconstructions and many more. The number of digital crimes equally increases as the new technologies, i.e. digital devices and internet, increases. As a result, we are becoming the soft target for various types of cybercrimes and digital attacks.

The Digital Forensic Research Workshop (DFRWS) has defined digital forensics (DF) as "The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations".

Todays, DF demands are increasingly important. DF investigation procedures help to capture important information from the compromised device. Nowadays, businesses deeply depend on the digital devices and on the Internet. Capturing the indispensable evidences from these devices is equally important. Digital evidence should be gathered from the system to support or deny some reasoning an investigator may have about the incident.

It is important to know that how to recover digital evidences which can be interested for investigators. However, current human power and other available resources are not enough to fully investigate the digital crimes on digital devices. Further, existing digital investigation procedures and practices require huge interaction with humans; as a result it slows down the process with the pace digital crimes are committed.

In this chapter, we have thoroughly discussed the current advancement of machine learning forensics (MLF) in digital forensic investigation (DFI). We present the latest surveys in this field and give critique comparisons of these approaches.

1.1 Historical perspective of digital forensic investigations

Digital forensic or computer forensic is first presented by 1970 [1]. In the first investigation, the financial fraud is proven from the suspect's computer. The first prosecuted computer crime was reported in 1996. The computer crime is defined as when the computer is the major effect for offense and facilitates the tool to

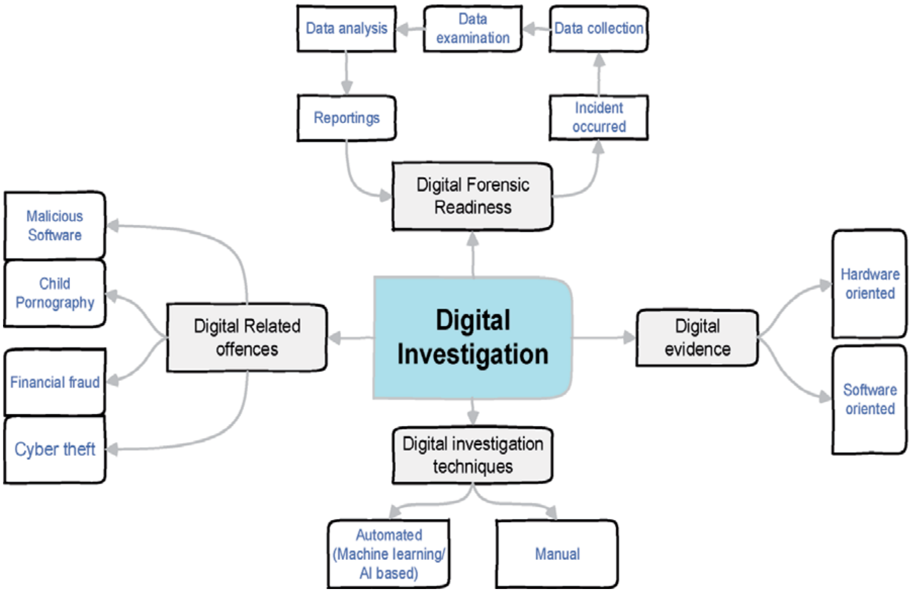


Figure 1.
Taxonomy of digital investigations.

commission a crime [2]. The first prosecuted computer crime was reported in Texas, USA, in 1996 [3] and resulted in a 5-year sentence. In 1990, computer-based digital crimes started to grow with the increasing popularity of the computers and the Internet. The computer forensic is developed as the independent field in the late 1990s and in the early 2000s. The CSI surveys report that almost 46% among the respondents were affected by some kind of computer crimes [4]. The 2010 Gallup surveys reports that 11% of the American adult become victim of computer- or Internet-related crimes in their homes. This ratio is 6–8% more than the last 7 years. A survey conducted by "Australian Company Crime Survey" [5], estimated that A$ 2,000,000 financial fraud and information breaches occurs in 2006. Company Crime Survey, its estimated A$ 2,000,000 financial fraud and information breaches in lost revenue. The term digital forensic is used nowadays with the advent of new digital devices with increasing number of frequency of use for investigation purposes (**Figure 1**).

2. Artificial intelligence (AI), machine learning (ML) and deep learning

It's important to examine how actually AI, ML and deep learning (DL) methods can help in solving the problems of DF and how these methods differentiate with each other's.

a. Artificial intelligence

AI is the science of making things smart or the capability of the machines (e.g. visual recognition, NLP, etc.) to perform human tasks. The important point is that AI is not machine learning or smart things. AI can be viewed as the things that can carry the human tasks and make these tasks easy. The AI technology is increasing day by day, and its enormous use also significantly increases the number of malicious activities.

Artificial intelligence programs are called intelligent agent. Intelligent agents are used to interact with the environment. The agent uses the technique to identify the environments through its sensors, and then it can take the action to affect the state through its sensors.

The important aspects in the AI technologies are how the sensors are used to collect the data and how it maps to the actuators; this is how the functions within the agents can perform these consequences. The ultimate goal of the AI is to develop the machine that acts just like humans. This task can be accomplished by only using the learning algorithms to which it is aimed to try to make a sketch of the human brain learnings. AI technologies give very good advantages and have a bright future ahead. However, these technologies are also unavoidably used for execution of some serious crimes that can be dangerous for people.

b. Machine learning

ML is one of the approaches of AI that uses a system that can be learned by itself from experience. It is not used for only AI purposes such as copying human behaviour but also needs to reduce the human efforts and time spent to perform the difficult and even the simple tasks. ML can be viewed as a system that can learn from experience and examples rather than from programming. Thus, if the system learns constantly and makes a decision based on the data rather than programming, then it's called ML. ML is developed as a new technology to provide new functionalities for computers and is used for industry and science. There are many autonomous solutions based on ML for medical science, robotics, engineering and so on.

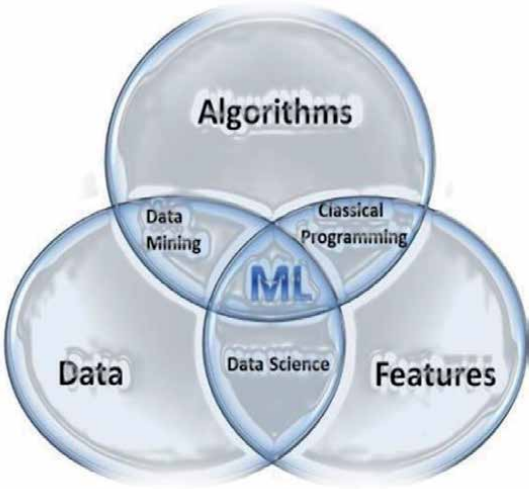


Figure 2.
Machine learning essentials.

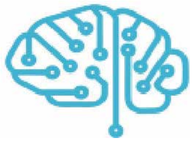
Artificial intelligence	Machine learning	Deep learning
		
Ability of a machine to imitate intelligent human behaviour	Application of AI that allows a system to automatically learn and improve from experience	Application of ML that uses complex algorithms and deep neural to train a model
Originated around the 1950s	Originated around the 1960s	Originated around the 1970s
Represents simulated intelligence in machines	Getting machines to make without being programmed	Process of using artificial neural networks to solve complex problems
Subsets of data science	Subset of AI and data science	Subset of ML, AI and data science
Building machines that are capable of thinking like humans	Make machines that can learn through previous experience to solve problems	To build neural networks that automatically discovers patterns for feature detection

Table 1.
Difference between artificial intelligence, machine learning and deep learning.

c. Deep learning

Deep learning combines the set of techniques used to implement ML methods to recognize patterns of patterns such as image recognition. First of all the system is used to identify the object edges, structure of the object, object type and then the object itself (**Figure 2**) (**Table 1**).

3. Approaches to machine learning forensics

Usually two main approaches are used to define the ML forensics, that is, inductive reasoning and deductive reasoning:

a. Inductive learning

Inductive reasoning is obtained from the general knowledge of specific information. The obtained knowledge is new and not truth preserving. That means the knowledge obtained can be invalidated from new information. There is no well-founded theory. In this area there are a large number of goals such as it is important to discover general concepts from a limited set of examples. The examples are called experience. The basis of this is to search for similar characteristics among examples. The methods used in these are based on the inductive learning.

b. Deductive learning

Deductive reasoning obtains the knowledge from well-established methods called logic. Deductive reasoning obtains from the knowledge by using well-established methods. The knowledge is not new. But it is implicit in the initial knowledge. New knowledge cannot invalidate the existing knowledge obtained and its basis on the mathematical logic.

3.1 Supervised, unsupervised and reinforcement ML

Supervised and unsupervised are the most commonly used techniques in ML algorithms.

a. Supervised and unsupervised

On the other hand, the reinforcement learning is complex and difficult to implement. Supervised learning is the most common type of ML paradigm. This type is easy to understand and implement. The data in this type is in the form of examples with labels. The data can be called as training data. The learning algorithms can be feed to these example-label pairs one by one. This allows the algorithms to predict the label for each example. Further, it provides the feedback whether this gives the right answer or not. In this type the model is first trained by using lots of training data (input and targets). This process is really fast and accurate. With the passage of time, the algorithms are able to learn in order to approximate the concrete nature of the relationship between examples and their labels. The trained supervised learning can see the totally new and never seen before data and predict the good label for it. Supervised learning is the most widely used and easiest to implement. Supervised learning is the most popular technique used for machine learning.

The unsupervised learning does not have a well-structured format. There are no targets for the training data. Therefore, the system does not know where to go. The system needs to understand itself from the given data. The unsupervised learning is the opposite of supervised learning. There are no labels in it. The algorithms are fed up with a lot of data, and the tool is given to understand the properties of the data. In this way, the task of the system is to learn to group, cluster and/or organize the data in the similar way as the human can organize the data. The unsupervised learning is much more interesting in a way that the overwhelming majority of data in this world is unlabelled. This type can make benefit of industries in a way that we have terabytes of unlabelled data, and organizing this data can be beneficial for the industry and potential profits for making it organized without minimal or no human effort (**Table 2**).

	Supervised learning	Unsupervised learning
Definition	Data set labeled with predefined classes	Data set labeled without predefined classes
Method	Data classification	Data clustering
Example	Support vector machine Decision tree	K-means clustering, ant clustering algorithms
Known attack detection	High	Low
Unknown attack detection	Low	High

Unsupervised learning is not easy and is not used as widely as supervised.

Table 2.
Supervised vs. unsupervised learning.

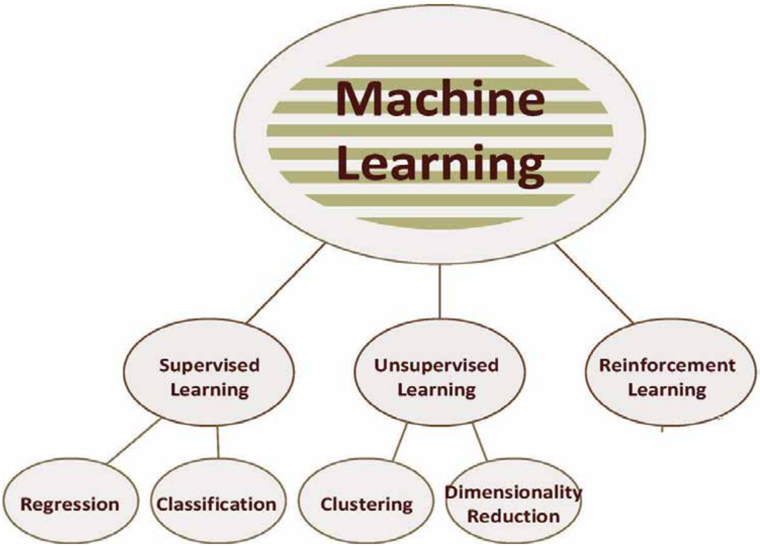


Figure 3.
Machine learning types.

b. Reinforcement learning

The reinforcement learning is totally different from both supervised and unsupervised ML. The relationship among supervised and unsupervised can be related with each other with the presence and absence of labels. However, the reinforcement learning learns from the mistakes. When deploying the reinforcement learning algorithms in any type of environment, it will make a lot of mistakes at the beginning. The signals to the algorithms are provided that can associate the good behaviour with positive signals and bad behaviour with negative label. The algorithms can reinforce algorithms to prefer good behaviour and bad behaviors. With the passage of time, the algorithm can learn to make fewer mistakes as it was initially (**Figure 3**).

3.2 Machine learning forensics for law enforcement, compliance and intelligence

Standardization is still a big challenge for DFI. The DI experts perform DI on the basis of their experience, the company's policies and basis on their previous experience. This is due to the lack of any universal standard for digital evidence collection. The law enforcement is continuously changing in this information technology age. The traditional crimes such as financial and commerce are also gaining the benefits of technology advancements and continuously upgrading with the latest development in the technology.

These days, law enforcement techniques are also changing.

DFI is a very common practice in law enforcements and commerce industry. The way in which the use of information technology is increasing by the government sectors, public and corporate agencies, has also increases the victimology of cyber-attacks through the internet.

4. Literature review

The work of [6] is one of the earliest efforts to make an application for expert systems for digital forensic to automate the analysis process. The expert system is used with decision tree in order to detect network anomalies automatically. The expert system is used to analyze the log files.

The Open Computer Forensic Architecture (OCFA) [7] is a well-organized forensic platform of automating the digital forensic tasks. This toll provides the scalability, modularity and openness in digital forensic process. This framework consists of different modules, and each module works independently on a specific file type in order for content extraction of the file for digital evidence. It creates the searchable index of text and metadata. It is a pluggable module that recursively processes the evidence according to the dispatching entity which decides which module needs to be invoked by seeing information in evidence. However, the OCFA follows the pre-extracted data and is not designed to search and recover files. The examination is done by an IT expert on the extracted data to generate indices for text and metadata.

Another effort is made by [8] of automating the disk forensic process. They name their tool “fiwalk” which is used to automate the processing of the data in order to assist the user for the development of the program which automatically processes disk images. This tool also integrates the command line tool of [9]. However, this toll only works for file system data only without any integration of AI techniques. Expert examiner tasks become easy by using this tool.

The research work of Hoelz et al. develops the MultiAgent Digital Investigation toolKit (MADIK) toolkit [9]. The tool provides the multiagent systems which helps the experts in computer forensic examinations. The authors apply the AI-based methods to the problem of digital forensics applications by assigning the tasks to each agent. Every agent has specialized in different tasks such as hashing, keyword search, Windows registry agent and so on. However this tool is not focused on building the new knowledge during investigations. It is used to learn from the previous investigations for any future investigation purposes. Moreover, this work cannot be used for nonexpert users.

The chapter [10] presents the machine learning-based digital triage model for selective pre-examination and statistical classification of digital data. This data can be deployed both on the crime scene and on digital forensic labs. The work is able to provide the quick actionable intelligence on the crime scene in time-critical systems, reduce the burden on forensic labs and protect suspect privacy when a huge amount of data is needed to be analyzed. As advantages the framework provides the minimum manual work and also produces measurable and reproducible error rate.

Existing methods for digital evidence extraction are not coherent to provide the readiness of process support with standardized integrated implementation system which provides guidance and technical knowledge to nonexpert investigators.

Paper title	Problems addressed	Methods used	Proposed solution	Implementation	Open problems
Building an intelligent assistant for digital forensics [11]	Supports investigations conducted by non-IT expert and expert investigators	Series of experiments comparing it with a human investigator as well as against standard benchmark disk images	Proposed AUDIT, an automated disk investigation toolkit	Systematically examine the disk in its totality based on its physical and logical structures	Seizure of an entire hard disk drive is a complex task
A Machine Learning-based Triage methodology for automated categorization of digital media [12]	Defines a list of crime-related features	Populates an input matrix and processes it with different machine learning mining schemes to come up with a device classification	Crime features extracted from available devices and forensic copies	Classified digital media using Bayes networks or support vector machines	Extract data in its raw form without the nature of the information
A machine learning-based approach to digital triage [13]	Identifies test objects allegedly used for exchanging child pornography material	Mobile handset classification on the basis of the 5MF technique	Multiclass categorization for classifying objects on the basis of owner's usage profile	Data corpus with binary categorization	Most files of forensic interest are not fragmented
Artificial intelligence applied to computer forensics [9]	Assists the computer forensic expert on its examinations	Set of rules and a knowledge base	MultiAgent Digital Investigation toolKit based on the experience of the expert	Six specialized intelligent agents implemented: HashSetAgent FilePathAgent FileSignatureAgent TimelineAgent WindowsRegistryAgent KeywordAgent	The method is very heavyweight to be practical
Automating disk forensic processing with SleuthKit, Xml and Python [8]	Automation to perform disk forensic	XML methods used to describe partitions and files on a hard drive or disk image	Creating special-purpose forensic tools	SleuthKit, XML and the Python programming language	Capturing every aspect of a live system is not feasible

Paper title	Problems addressed	Methods used	Proposed solution	Implementation	Open problems
Automated analysis for digital forensic science: Semantic integrity checking [6]	Automates data collection	Expert system with a decision tree	Predetermined invariant relationships between redundant digital objects to detect semantic incongruities	Collection of C programs and Perl scripts	
Android forensics: Automated data collection and reporting from a mobile device [14]	Broadcasts receiver, content observer and alarm	Forensic collection, local SQLite storage, HTTP transfer and clear local SQLite DB	Collects, stores and transfers forensically valuable Android data to a remote Web server without root privileges	DroidWatch is an automated system prototype composed of an Android application and an enterprise server	Architecture models of Android applications are complex and diverse in nature
An automated approach for digital forensic analysis of heterogeneous big data [15]	Understanding the relationships between artifacts	Metadata to solve the data volume problem, semantic web ontologies to solve the heterogeneous data sources	Automated identification and correlation	Artifacts to reduce the burden placed upon the investigator	Not given any particular implementation details
Data mining methods applied to a digital forensics task for supervised machine learning [16]	Glass identification in the context of multi-class supervised learning	Decision trees, Bayes classifiers, based on rules, artificial neural networks and based on nearest neighbors	Empirical overview of the performance with classifiers from different machine learning approaches	Uses two metrics like accuracy and Cohen's kappa for training and test stages	Abstraction errors can occur when representations of the system are not accurate
Data mining methods applied to a digital forensics task for supervised machine learning	Multi-class classification	Supervised machine learning techniques	Decision trees, Bayes classifiers, based on rules, artificial neural networks and based on nearest neighbour techniques	Nondeterministic algorithms	The algorithms implemented are complex in nature and system needs careful understanding of the extracted data
Android forensics: Automated data collection and reporting from a mobile device [14]	Enterprise monitoring system for Android smartphones	Comprehensive guide of data sets available for collection without elevated privileges	First open-source Android enterprise monitoring prototype	Continuously collect many data sets of interest to incident responders, security auditors, proactive security monitors and forensic investigators	Increasing interoperability among Android devices

Paper title	Problems addressed	Methods used	Proposed solution	Implementation	Open problems
Automated forensic analysis of mobile applications on Android devices [17]	Inter-component string propagation, string operations (e.g. append) and API invocation	Inter-component static analysis on Android APKs	Identifies how the information is stored by parsing SQL commands	Fordroid: builds control flow and data dependency graphs	Inter-component string propagation
Automated inference of past action instances in digital investigations [18]	Detects multiple instances of a user action	Signature-based methods	Integrating time into event reconstruction	Detected using signature-based methods during a postmortem digital forensic analysis	Aligning time stamps from different systems and analyzing complex events with incomplete time information
An automated timeline reconstruction approach for digital forensic investigations [19]	Extracts low-level events to a SQLite backing store	Pattern matching to automatically reconstruct high-level, human understandable events.	Automatically analyzed for patterns	Automatically reconstruct high-level events (e.g. connection of a USB stick) from this set of low-level events	Do not cover all aspects of forensic analysis between events
Automated event and social network extraction from digital evidence sources with ontological mapping [20]	High-level analysis based on low-level digital artifacts	Automatically derived "events" from the base forensic artifacts	Information fusion and homogenization techniques are used to reconstruct social networks	Standardized knowledge representations techniques and automated rule-based systems to encapsulate expert knowledge for forensic data	Validation of extracted ontologies and correctness of the data is a big issue

Table 3.
Comparison of literature surveys.

The lack of the automated intelligent systems for digital evidence extractions is another big issue. Further, digital evidence are difficult to handle and cannot be easily understandable even for experts. Extracting digital evidence from different storage media may require several layers of transformations (**Table 3**).

5. The significance of machine learning in digital forensic investigations

MLF is originating from AI to perform the huge amount of data, analyse the data to discover any criminal actions and risk and to segment the data to find criminal activity and behaviour. The intelligence systems which do not have any intelligent part cannot perform true learning capabilities and be a true one. DFI through ML is the latest trend to seize the potential of AI as leading security solutions capabilities.

ML behavioral analytics is the core part of modeling, profiling and prediction in medical, manufacturing, advertising and business intelligence and is recently used in law enforcement mechanism. In order to discover the criminal behaviour, MLF uses the wireless or wired networks via web or cloud computing. Thus MLF aims are to provide the new knowledge and skills and provide organized knowledge structure in order to produce progressive improvements in its own performance.

Originating from AI, ML algorithms can be used to analyze the huge amount of data to identify the risk, segment the data and detect criminal behaviour. ML algorithms enable the investigators to interrogate the vast scattered data sets which are placed in social and wired networks and web or cloud computing. In essence, ML algorithms contain the pattern recognition software that are used to analyse huge amount of data which are used to predict some behaviour. ML algorithms seek to learn from historical perspectives which are then used to predict future behaviour. MLF gains the capability to recognize the patterns of criminal activities through ML algorithms, in order to learn from the historical data about when and where the crime will take place. The malicious activities from extracted data set can be from burglaries, money laundering or intrusion attacks. This task can be achieved by formalizing and analyzing the servers, suspect's devices, wireless devices, the Internet and other kinds of data for visualization, link association, segmentation and predicting criminal activities. Nowadays, the industry is facing more advance cyber threats that cannot be tracked though traditional security measures. Attackers have designed more sophisticated ways to attacks on the system and become complicated over time. System administrator would not be able to detect these attacks each time. On the other hand, human expertise and competences have some limits, and this leads to the fact that industry is lacking in poor speed of incident occurrence, longer delay in detection and prevention of cyber threats and takes more advanced expertise to remove these cyber threats. Therefore, developing more advance machine learning models may help to prevent and protect form these cyber threats. Nowadays, there are many automated software available that can help the human to perform complicated and scientific tasks. In the next step, these automated tools need to be more advanced and should have the capability of AI and ML techniques.

6. Discussion and future prospects

From literature survey, it has been observed that there are many challenges which can be faced by the forensic experts when performing the test.

First of all there is an ultra-exponential growth in the data due to the inexpensive storage devices such as hard drives, CD, USB stick and so on. This makes it almost impossible for the individuals to perform the forensic in a short period of time.

Consequently, it is almost impossible for the forensic experts to perform the proper data analysis of each machine individually and also perform the cross-check on each machine's process. That limits the capability of the human works. In this line of reasoning, a huge amount of data needs to be sent to laboratory for forensic purposes with limited time and available resources. In a real-time digital forensic investigation, it is very difficult to determine in early stages which evidence is more important and relevant for investigating the crime, as an example, if we consider the cybercafé or a network of computers where several computers share the same IP address.

On the other hand, the intelligent tools are the main part of the MLF. However, these tools also show the problem for investigation in the pre-analysis phase. For that reason the lack-ness in the collection of large amount of data from distributed machines is need to be examined. Some of the existing tools are not helpful in solving the problem and even increases the time of investigation. The need is to make more intelligent methods and tools so that the automatic investigation of the suspects machines or malicious activity can be analyzed and determined in accurate time. The data can be stored and placed in any place for destructive purposes. Therefore, MLF techniques are the best sources for storing, evaluating and using this data in a productive way to anticipate and harmful activities. MLF methods can perform the meta-analysis on the meta-knowledge from different sources, and it can simplify the complex tasks into understandable and manageable data formats in a short period of time. MLF can provide the well-formed repository that can contain the well-sanitized data of digital investigation with well-known properties and results.

- Machine learning forensics solutions should:
    - Have data availability to support modeling.
    - Address well-scoped problems and methodology.
    - Explain well the reasoning process.
    - Formally structure the representation of knowledge.
    - Have well-organized performance evaluation.
    - Integrate with current architecture, tools and applications.

Author details

Salman Iqbal[1]* and Soltan Abed Alharbi[2]

1 Department of Computer Science, COMSATS University Islamabad, Vehari, Pakistan

2 Department of Electrical and Computer Engineering, University of Jeddah, Saudi Arabia

*Address all correspondence to: simbwp@gmail.com

References

[1] Pollitt M. A history of digital forensics. In: IFIP International Conference on Digital Forensics. Springer, Berlin, Heidelberg: Springer; 2010. pp. 3-15

[2] Raghavan S. Digital forensic research: Current state of the art. CSI Transactions on ICT. 2013;**1**(1):91-114

[3] Dierks MP. Computer network abuse. Harvard Journal of Law & Technology. 1992;**6**:307

[4] Richardson R, Director C. CSI computer crime and security survey. Computer Security Institute. 2008;**1**:1-30

[5] A.C.E.R.T.A. 2006 Australian Computer Crime and Security Survey. AusCERT & Australian High Tech Crime Center (AHTCC); November 23, 2006. Available from: http://www.auscert.org.au/render.html?it=2001

[6] Stallard T, Levitt K. Automated analysis for digital forensic science: Semantic integrity checking. In: 19th Annual Computer Security Applications Conference, 2003. Proceedings. IEEE; 2003

[7] Vermaas O, Simons J, Meijer R. Open computer forensic architecture a way to process terabytes of forensic disk images. In: Open Source Software for Digital Forensics. Boston, MA: Springer; 2010. pp. 45-67

[8] Garfinkel SL. Automating disk forensic processing with SleuthKit, XML and python. In: 2009 Fourth International IEEE Workshop on Systematic Approaches to Digital Forensic Engineering. IEEE; 2009

[9] Hoelz BW, Ralha CG, Geeverghese R. Artificial intelligence applied to computer forensics. In: Proceedings of the 2009 ACM Symposium on Applied Computing. ACM; 2009

[10] Fizaine J, Clarke N. A crime depended automated search and engine for digital forensics. Advances in Communications, Computing, Networks and Security. 2013;**10**:73

[11] Karabiyik U. Building an Intelligent Assistant for Digital Forensics. 2015

[12] Marturana F, Tacconi S. A machine learning-based triage methodology for automated categorization of digital media. Digital Investigation. 2013;**10**(2):193-204

[13] Marturana F, Tacconi S. A machine learning-based approach to digital triage. Methodology for Automated Categorization of Digital Media. In Digital Investigation. Elsevier; 2013;**10**(2):193-204

[14] Grover J. Android forensics: Automated data collection and reporting from a mobile device. Digital Investigation. 2013;**10**:S12-S20

[15] Mohammed H, Clarke N, Li F. An Automated Approach for Digital Forensic Analysis of Heterogeneous Big Data. The Journal of Digital Forensics, Security and Law. 2016

[16] Tallón-Ballesteros AJ, Riquelme JC. Data mining methods applied to a digital forensics task for supervised machine learning. In: Computational Intelligence in Digital Forensics: Forensic Investigation and Applications. Springer; 2014. pp. 413-428

[17] Lin X et al. Automated forensic analysis of mobile applications on android devices. Digital Investigation. 2018;**26**:S59-S66

[18] James JI, Gladyshev P. Automated inference of past action instances in digital investigations. International Journal of Information Security. 2015;**14**(3):249-261

[19] Hargreaves C, Patterson J. An automated timeline reconstruction approach for digital forensic investigations. Digital Investigation. 2012;**9**:S69-S79

[20] Turnbull B, Randhawa S. Automated event and social network extraction from digital evidence sources with ontological mapping. Digital Investigation. 2015;**13**:94-106

Chapter 2

Cybersecurity Risk Analysis of Industrial Automation Systems on the Basis of Cognitive Modeling Technology

Vladimir I. Vasilyev, Alexey M. Vulfin and Liliya R. Chernyakhovskaya

Abstract

The issues of procuring the cybersecurity of modern industrial systems and networks acquire special urgency because of imperfection of their protection tools and presence of vulnerabilities. International standards ISA/IEC 62443 offer the system risk-oriented approach to solve the tasks of providing the security of industrial control systems (ICS) at all stages of life cycle. But in view of high uncertainty and complexity of procedure of formalizing the factors affecting the final indices of system security, the problem of cybersecurity risk assessment remains open and requires applying new approaches based on the technology of data mining and cognitive modeling. Cognitive modeling of risk assessment using fuzzy grey cognitive maps (FGCM) allows us to take into account the uncertainty factor arising in the process of vulnerability probability assessment for each of security nodes. The interval estimates of FGCM connection weights can reflect the scatter of expert group opinions that allows us to take into account more completely the data available for risk analysis. The main stages of ICS security assessment with use of FGCM are analyzed in the chapter on the example of distributed industrial automation network. The recommendations concerning the choice of the necessary countermeasures improving the level of network security in the conditions of possible external and internal threats are considered.

Keywords: fuzzy grey cognitive map, cybersecurity risk analysis, industrial automation systems, cognitive modeling, integrity control model automation system

1. Introduction

Digital economy, cyber-physical objects, cyberspace, and Internet of Things are concepts that have firmly entered our lives in recent years. As a part of industrial revolution “Industry 4.0,” the face of modern industrial enterprises, which actively use the transition to unmanned production technologies, the integration of information technologies into the most complex production processes, has dramatically changed. In this case, a distinctive feature of production is the close connection of technological networks with the corporate network, as it is necessary both for production management and for administration of industrial networks and systems. Modern technological networks, as a rule, have direct access to the Internet, for example, for maintenance and technical support of industrial automation systems by employees of organizations—contractors. Also, computers of contractors, developers, integrators, and system/network administrators connected to the technological network of the service company from the outside often have free access to the Internet.

Under such conditions, the problem of ensuring the security (cybersecurity) of industrial automation and control systems (IACS) sharply increases. In corporate networks, the object of protection is information and the problem of ensuring the confidentiality of information is primarily addressed. However, in the case of industrial automated control systems, the object of protection is already technological processes (TP), and not ensuring the confidentiality of information comes to the fore, but first of all ensuring the continuity and integrity of the TP itself. Speaking of IACS cybersecurity, the so-called digital attacks (cyber-attacks) are primarily considered associated with exposure to IACS through the control and monitoring devices—controllers, data acquisition and transmission devices, SCADA servers, workstations, telecom equipment, communication lines, etc.

The severity and relevance of IACS cybersecurity problem are confirmed by statistics of recent years, showing a sharp increase in the number of the targeted attacks on industrial networks and systems, as well as an increase in the scale of consequences of these attacks. A vivid example of a large-scale cyber-attack that hit a lot of companies around the world from May 12 to May 15, 2017 is the attack of a network worm—the coder WannaCry [1]. Among the victims of this well-coordinated attack were companies engaged in various types of production, oil refineries, urban infrastructure facilities, and distribution power grids.

In May 2018, VPNFilter malware, which infected at least 500,000 routers and data storage devices in 54 countries around the world, was detected. The purpose of this software is to steal credentials, detect industrial SCADA equipment, and carry out various attacks using infected devices in the botnets. June 2018 was marked by a large-scale cyber-attack on telecommunications companies, communication satellite operators, and defense contractors in the United States and Southeast Asia. During the attack, the attackers infected computers used for managing the communication satellites and collecting geoposition data. According to experts' opinions, the purpose of the cyber-attack was espionage and data interception from civilian and military communication channels. In total, according to Kaspersky Lab, the share of attacked IACS computers in the world in 2018 increased by 3.2% compared with 2017 and amounted to 47.2% [2].

Considering the seriousness of the current situation and the need to take urgent measures, the international community and information security experts are concerned about finding effective ways to solve the problem of ensuring the security of industrial automated systems. For instance, the European Commission has developed the European Program for Critical Infrastructure Protection. Several international standards for ensuring the cybersecurity of automated process control systems have been proposed and effectively used in world practice, such as NERC Critical Infrastructure Protection, NIST SP 800-82 Guide to Industrial Control Systems Security, ISA/IEC 62443 Industrial Automation and Control Systems Security [3, 4].

The basis of the requirements presented by the ISA/IEC 62443 standards series for ensuring the IACS security is a risk-oriented approach. In accordance with this approach, designing of a management system for a protected IACS involves the following stages:

- high-level (preliminary) risk assessment of cyber-attacks effects;

- building a reference model of IACS as the protection object, describing the classification of main activities types, technological process, automatic control systems, and other assets;

- building an asset model, describing the hierarchy of main objects and assets of IACS, their interaction with networks, key divisions, etc.;

- building a reference architecture model, reflecting all basic elements of IACS, telecommunication equipment, communication lines, etc.;

- building a zone and conduct model, dividing the protected object into separate security zones;

- detailed risk analysis for each selected zone; and

- determination of the current security level for each zone and requirements to ensure the target security level of the zone, implemented by the choice of appropriate protection measures.

At the same time, the "bottleneck" of the above normative documents regulating the issues of ensuring IACS cybersecurity is the absence of formalized methods for detailed risk assessment. As the volume of statistical data, development of mathematical models of risk, threats, and security incidents increase, it becomes topical to develop methods and algorithms for quantitative risk assessment, ensuring the possibility of a reasonable choice of IACS devices and the necessary countermeasures both within individual security zones and ensuring the required cybersecurity level of IACS as a whole.

A promising way to solve this problem is the use of technology of cognitive modeling, based on construction and analysis of fuzzy grey cognitive maps (FGCM), which has been widely used in recent years [5–10]. Fuzzy grey (interval) cognitive maps are considered to be a good extension of fuzzy cognitive maps (FCM) family, since they are better suited to experts representations, have a greater interpretability and provide more degrees of freedom to the decision making person on the basis of modeling results.

Brief information concerning the "grey" system, the "grey" number, and the "grey" variable is presented below, and the mathematical apparatus of FGCM is considered. Then, on the example of solving the problem of ensuring the integrity of telemetric information in IACS, the technique of assessing the cybersecurity risks with use of FGCM is discussed. In the end of the chapter, the conclusions are drawn and the list of references is given.

Let us note one important circumstance. When considering below a specific example of AIS risk assessment using FGCM (Section 2), an approach based on decomposition of the original (integrated) FGCM by disclosing (detailing) the content of its concepts is used, resulting in the set of interconnected local FGCM that characterize certain aspects of AIS risks assessment procedure associated with the features of its subsystems. In ideological plan, this approach is based on the FCM decomposition theory and the algebra of FCM causal transformations proposed in [11, 12]. However, the main difference between the approach described in [12] and our approach is that in [12] the detailed FCM system of a large size comes out as the original FCM, which reduces to a simpler (quotient) FCM by using the operations proposed by the authors. Each concept of this quotient FCM accumulates information on the state of several similar concepts of the original FCM, thus aggregating the corresponding concepts. In our case, on the contrary, the original FGCM has a small dimension, the number of forming its basic concepts corresponds to the number of basic subsystems of the system under study, and the decomposition of FGCM implies a representation of each concept of the original FGCM in the form of independent (local) FGCM, describing the behavior of this concept.

2. Theoretical foundations of building FGCM

The basis of building FGCM is the Grey Systems Theory, proposed in 1989 by Deng [10]. Within the framework of this theory, objects and systems with high uncertainty, represented by small samples of incomplete and inaccurate data, are studied. Depending on the character of the available information, the studied systems are divided into three types:

- "white" systems (the internal structure and the properties of the system are completely known);
- "grey" systems (partial information about the system is known); and
- "black" systems (the internal structure and the properties of the system are completely unknown).

In accordance with the terminology of the grey systems theory, a fuzzy grey cognitive map is a cognitive model of a system in the form of a directed graph defined with use of the following set:

$$\mathrm{FGCM} = \langle C, F, W \rangle, \tag{1}$$

where $C = \{C_i\}$ is the set of concepts (vertices of the graph), $(i = 1, 2, \ldots, n)$; $F = \{F_{ij}\}$ is the set of connections between concepts (arcs of the graph); and $W = \{W_{ij}\}$ is the set of the relationships between the concepts determining the weights of these connections, $(i,j) \in \Omega$. Here, $\Omega = \{(i_1,j_1), (i_2,j_2), \ldots, (i_L,j_L)\}$ is the set of the pairs of adjacent (interconnected) vertices indices, $L \leq n(n-1)$.

In contrast to the traditional FCM representation, the weights of FGCM connections are set with the use of "grey" (interval) numbers $\otimes W_{ij}$, defined as

$$\otimes W_{ij} \in \left[\underline{W_{ij}}, \overline{W_{ij}}\right], \text{where } \underline{W_{ij}} < \overline{W_{ij}}, \left[\underline{W_{ij}}, \overline{W_{ij}}\right] \in [-1, 1], \tag{2}$$

where $\underline{W_{ij}}$ and $\overline{W_{ij}}$ are, respectively, the lower and the upper boundaries of the grey number. So, the weight of connection between i-th and$j$-th concepts $(C_i \to C_j)$ can take any value within the given range of change $\left[\underline{W_{ij}}, \overline{W_{ij}}\right] \in [-1, 1]$. In the particular case, when $\underline{W_{ij}} = \overline{W_{ij}}$, we get $\otimes W_{ij} \in \left[\underline{W_{ij}}, \underline{W_{ij}}\right]$—a "white" (crisp, usual) number.

It is assumed that the change of the concepts state in time is described by equations

$$\otimes X_i(k+1) = f\left(\otimes X_i(k) + \sum_{\substack{j=1 \\ (j \neq i)}}^{n} \otimes W_{ji} \otimes X_j(k)\right), (i = 1, 2, \ldots, n), \qquad (3)$$

where $\otimes X_i(k)$ is the "grey" (interval) variable of the i-th concept C_i state, the values of which at each time instant $k = 0, 1, 2, \ldots$ belong to some interval $\left[\underline{X_i}(k), \overline{X_i}(k)\right]$; f is the activation function of the i-th concept, mapping the argument values into the interval $[-1, 1]$. The activation function $f(\bullet)$, as a rule, is accepted in the following form:

a. linear function with saturation:

$$f(x) = \begin{cases} x, \text{if} |x| \leq 1; \\ \text{sgn } x, \text{if} |x| > 1, \end{cases} \qquad (4)$$

b. bipolar sigmoid (hyperbolic tangent):

$$f(x) = \frac{(1 - e^{-x})}{(1 + e^{-x})} = \text{th}\left(\frac{x}{2}\right); \qquad (5)$$

c. unipolar sigmoid:

$$f(x) = 1/(1 + e^{-x}). \qquad (6)$$

To solve the system of equations (Eq. (3)), it is required to set the initial conditions $\otimes X_i(0)$, which also should be considered as the grey numbers $\otimes X_i(0) \in \left[\underline{X_i}(0), \overline{X_i}(0)\right]$. Most interesting is usually to obtain the equilibrium (steady state) solution, which is a grey vector $\lim\limits_{k \to \infty} [\otimes X_i(k)] = \otimes X^* \in \left[\underline{X^*}, \overline{X^*}\right]$ or a limit cycle (strange attractor).

To determine the stability of the steady-state solution $\otimes X^*$, one can use the theorem [12], according to which the only equilibrium (steady state) solution of equation (3) ("the fixed point") exists if and only if

$$\left(\sum_{i,j=1}^{n} \overline{W}_{ij}^2\right)^{\frac{1}{2}} < H, \qquad (7)$$

where the value of the positive constant H depends on the choice of activation function of the concepts: $H = 1$ for function (Eq. (4)); $H = 2$ for function (Eq. (5)); and $H = 4$ for function (Eq. (6)). In the case of negative connection, i.e., for $\underline{W_{ij}} < \overline{W_{ij}} < 0$, we also put in (Eq. (7)) the upper boundary $\overline{W_{ij}}$ of the grey number $\otimes W_{ij}$.

More detailed information on FGCM construction and their learning algorithms can be found in [5, 6, 9].

3. Risk assessment of IACS cybersecurity

Let us consider the task of assessment of IACS risk on the example of the automated system for collecting, storing, and processing the telemetric information (TMI) of the aviation equipment manufacturer. The current information on the state parameters of on-board aviation systems is continuously collected during the entire period of their operation by the ground services of technical maintenance. The detailed analysis of this information allows the subsequent making the right management decisions on the further operation and modification of on-board aviation systems. Therefore, the task of ensuring the integrity of the mentioned telemetric information under the conditions of possible impact of external and internal threats undoubtedly takes on particular significance.

The generalized structure of the studied territorially distributed automated information system (AIS), providing the collection, storage, and processing of TMI, is presented in **Figure 1**.

As the parts of AIS, the following subsystems (zones), combined according to the principle of the unity of functions performed and security requirements for their implementation, are identified:

1. The subsystem for collecting and storing the primary data at the service stations (Zone 1), which includes:

   Element 1—the client part of the SCADA system Web-base;

   Element 2—the server part of the SCADA system Web-base;

   Element 3—OPC UA client;

   Element 4—the temporary storage for accommodating the operative telemetry data accumulated at the object;

   Element 5—the server part of the accumulated data transmission to the storage of the aviation equipment manufacturer;

2. The core of the corporate information network (CIN) of the enterprise-manufacturer (Zone 2), where:

   Element 6—the client part for providing access to the server of the service station transferring the accumulated operational data of TMI to the enterprise-manufacturer’s storage;

   Element 8—the workstations of administrator and service personnel of the CIN core of the enterprise-manufacturer;

3. TMI storage subsystem with fault tolerance functions (Zone 3), where:

   Element 7—the node of access to TMI data storage at the enterprise-manufacturer;

4. TMI data processing subsystem with the use of a hierarchy of mathematical models of aviation equipment (Zone 4);

5. Subsystem of support and implementation of business processes of the enterprise-manufacturer (Zone 5).

The corresponding subsystems (security zones) are interconnected (see **Figure 1**) with the aid of telecommunication channels (conducts).

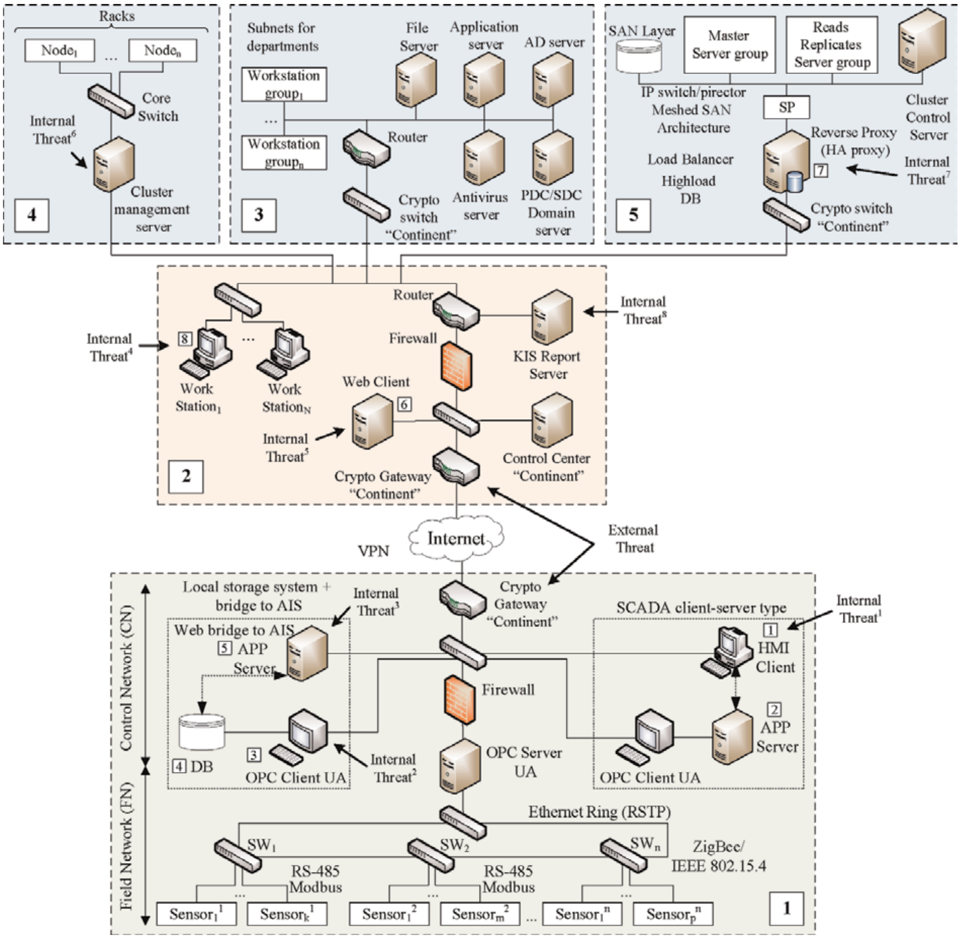


Figure 1.
The generalized structure of territorially distributed automated information system for the collection, storage, and processing of TMI. The corresponding subsystems (security zones) are interconnected with the aid of telecommunication channels (conducts).

Using FGCM as a tool for cognitive modeling, let us turn to the task of analyzing the risks associated with ensuring the TMI integrity in AIS considered above, taking into account the impact of possible external and internal threats to the system. The original (integrated) FGCM for assessing the risks of AIS, serving in this case as the AIS cognitive model of initial approximation (zero decomposition level), is presented in **Figure 2**.

The following descriptions are used in **Figure 2**: superscript ("*") denotes the affiliation of the concept C_p^* to integrated FGCM and subscript (p) denotes the number of the concept (**Table 1**).

The presence of the grey connection weights $\otimes \tilde{W}_{ij}$ indicates an uncertainty in the assessment of the mutual influence of main risk factors. The state variables of concepts $\otimes X_{T_1}^*$, $\otimes X_{T_2}^*$, $\otimes X_i$, $(i = 1, 2, ..., 5)$, $\otimes X_R^*$ represent the probabilities of occurrence of the enumerated events corresponding to the concepts $C_{T_1}^*$, $C_{T_2}^*$, C_1^*, ..., C_5^*, C_R^*. Let us note that in this case we mean so-called subjective probabilities, reflecting the expert's point of view on the possibility of an event occurrence [13]. Taking into account that each of these events is a complex event consisting of a chain of consecutive elementary events, it is reasonable to decompose FGCM of AIS

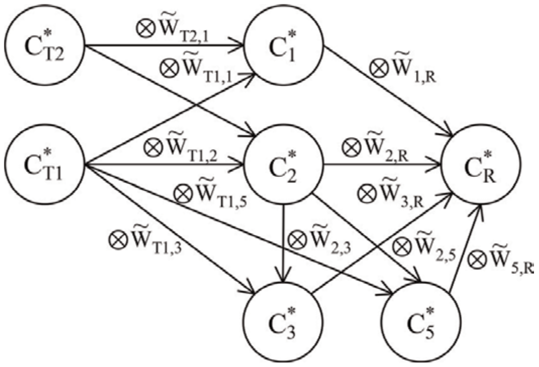

C^*_{T2}
$\otimes\widetilde{W}_{T2,1}$
C^*_1
$\otimes\widetilde{W}_{T1,1}$
$\otimes\widetilde{W}_{1,R}$
C^*_{T1}
$\otimes\widetilde{W}_{T1,2}$
C^*_2
$\otimes\widetilde{W}_{2,R}$
C^*_R
$\otimes\widetilde{W}_{T1,5}$
$\otimes\widetilde{W}_{3,R}$
$\otimes\widetilde{W}_{T1,3}$
$\otimes\widetilde{W}_{2,3}$
$\otimes\widetilde{W}_{2,5}$
$\otimes\widetilde{W}_{5,R}$
C^*_3
C^*_5

Figure 2.
Integrated FGCM for AIS risk assessment. $\otimes\tilde{W}_{ij}$*—the grey connection weights indicate an uncertainty in the assessment of the mutual influence of main risk factors and* C^**—concepts.*

Concept	Concept name
$C^*_{T_1}$	Internal threat to TMI integrity (e.g., due to failures or erroneous actions of staff)
$C^*_{T_2}$	External threat to TMI integrity (e.g., due to attempts of unauthorized access from outside to information)
C^*_1	Modification of TMI data in Zone 1
C^*_2	Modification of TMI data in Zone 2
C^*_3	Modification of TMI data in Zone 3
C^*_5	Modification of TMI data in Zone 5
C^*_R	Potential damage caused by violation of TMI integrity in AIS

Table 1.
List of the concepts of the integrated FGCM.

shown in **Figure 2** as the set of FGCMs for separate concepts (AIS security zones containing targets objects for attack to TMI).

The first decomposition level of the original (integrated) FGCM is presented in **Figure 3**.

The following designations of the concepts are used in **Figure 3**: the superscript (q) of C^q_p indicates the belongings to the concept C_q of the integrated FGCM; and the subscript (p) is the number of the concepts in the FGCM of the first level of decomposition (**Table 2**).

Figure 4 shows the further decomposition level (the second level) for the concept C^*_1, allowing to make clearer the impact of the threats on the considered target concept.

On the scheme, the following designations of the concepts of FGCM second-level decomposition are used: the superscript (q) of the $C^{q,p}_r$ concept is the number of the concept (the parent concept of the zero decomposition level) of the original FGCM whose decomposition includes this element, the superscript index p is the number of the parent concept of the first level of decomposition, the subscript (r) is the number of the concept of the current level (**Table 3**).

The further decomposition of the third level allows us to go to the detailed FGCM, which allows us to take into account the influence of individual vulnerabilities on the potential violation of TMI integrity in the intermediate information processing elements.

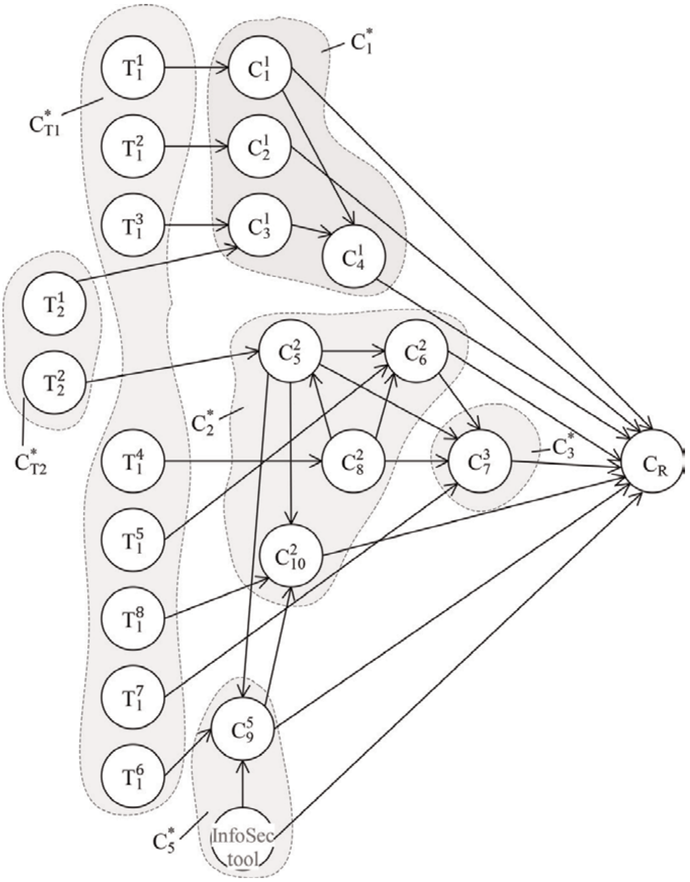


Figure 3.
The first level of FGCM decomposition to assess the AIS risks.

As for the concept $C_1^{1,1}$, characterizing the possibility to run in the browser of the client part of SCADA system on the base on Web technology (Zone 1), the corresponding decomposition can be represented as FGCM in **Figure 5**.

Here, the numbers 1–5 denote the following concepts:

1. the exploitation of the vulnerability of OS authorization system;

2. the exploitation of the vulnerability of SCADA Web client;

3. the exploitation of the vulnerability of OS browser for launching the client part of SCADA;

4. the exploitation of the vulnerability of access to OS memory;

5. the exploitation of the vulnerability of OPC UA client authorization system.

Similarly, it is possible to decompose the other concepts of original FGCM for the second decomposition level of Zone 1 presented in **Figure 4** (**Figures 6**–**9**, **Tables 4–6**). The corresponding FGCM, revealing the content of the concept $C_2^{1,1}$ (Zone 2), is shown in **Figure 6**.

Concept	Concept name	Parent concept
$T_1^1 - T_1^8$	Internal threats to the integrity of TMI (concept T_1^* decomposition on the block diagram of AIS, **Figure 1**, i.e., the points of potential realization of the threat to TMI integrity by the internal subject of the system)	T_1^*
T_2^1, T_2^2	External threats to TMI integrity (concept T_2^* decomposition)	T_2^*
C_1^1	Access to TMI in the client-server SCADA Web-base before adding to the database of TMI operational storage	C_1^* (Zone 1)
C_2^1	Access to the database of operative TMI data storage	
C_3^1	Access to the network equipment	
C_4^1	Access to the module of Web server sending TMI data in the long-term storage of the enterprise-manufacturer	
C_5^2	Access to the network infrastructure	C_2^* (Zone 2)
C_6^2	Access to the Web client module that implements receiving TMI at the enterprise-manufacturer from remote service stations	
C_8^2	Unauthorized access to workstation (node 8 in **Figure 1**) of the core of CIN of the enterprise-manufacturer	
C_{10}^2	Access to the server of equipment status reports generated for users of Zone 4	
C_7^3	Access to TMI in the long-term storage	C_3^* (Zone 3)
C_9^5	Access to computing cluster management server of Zone 5	C_5^* (Zone 5)
IST^5	TMI integrity control model	

Table 2.
List of the first level decomposition concepts of the FGCM.

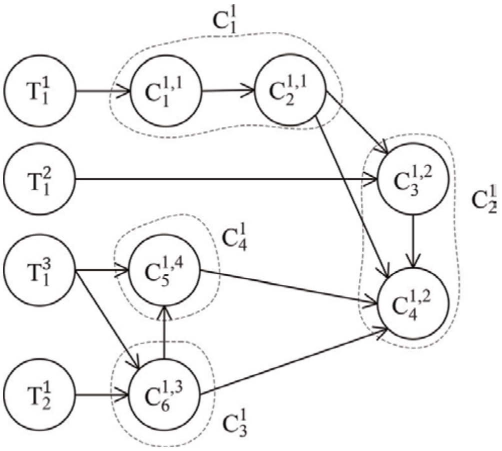


Figure 4.
The second level of FGCM decomposition for assessing AIS risk in Zone 1.

Consider the numerical example of risk assessment for the concept $C_1^{1,1}$ (**Figure 5**).

Let us assume that while choosing the grey values of FGCM weights, it is necessary to focus on a certain fuzzy scale, which determines the strength of the connections between different concepts (see, e.g., **Table 7**).

Let us further assume that the expert estimated the values of FGCM connections weights in **Figure 5** as follows (**Table 8**).

Concept	Concept name	Parent concept
$C_1^{1,1}$	Access to HMI client SCADA	C_1^1
$C_2^{1,1}$	Access to operative TMI data on the client-server part of the SCADA before entering in the operative storage	
$C_3^{1,2}$	Access to the client to interact with the OPC UA server	C_2^1
$C_4^{1,2}$	Access to the database of operative TMI storage data	

Table 3.
List of second-level decomposition concepts for Zone 1.

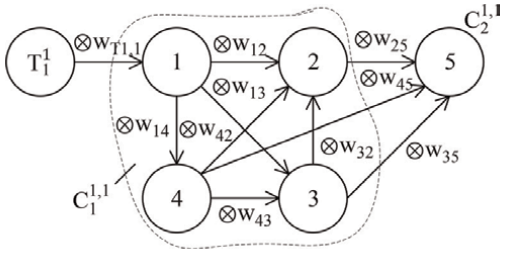


Figure 5.
The third level of FGCM decomposition—the concept $C_1^{1,1}$.

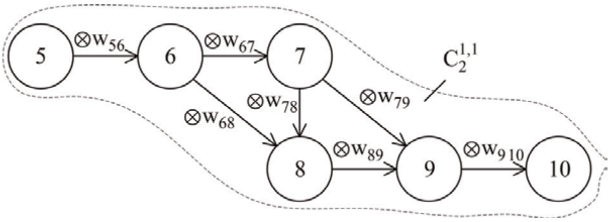


Figure 6.
Decomposition of the concept $C_2^{1,1}$ of FGCM for AIS risk assessment (Zone 1)

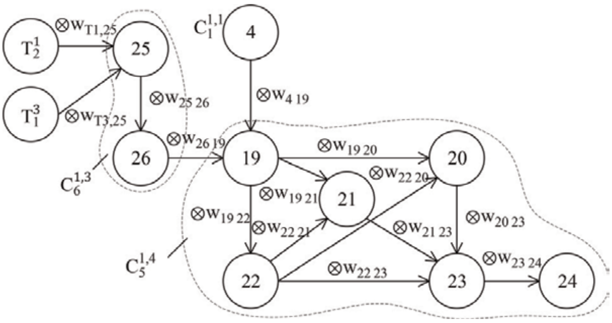


Figure 7.
Decomposition of the concepts $C_6^{1,3}$ and $C_5^{1,4}$ of the second level of FGCM decomposition

Let us take a bipolar sigmoid (5) here as an activation function $f(\bullet)$ for the concepts 1–5. Checking condition (7) for the data presented in **Table 2** shows that

$$\left(\sum_{i,j=1}^{5} \overline{W}_{ij}^{2}\right)^{\frac{1}{2}} = \sqrt{2,76} = 1.66 < 2, \tag{8}$$

i.e., the steady-states of FGCM will be stable.

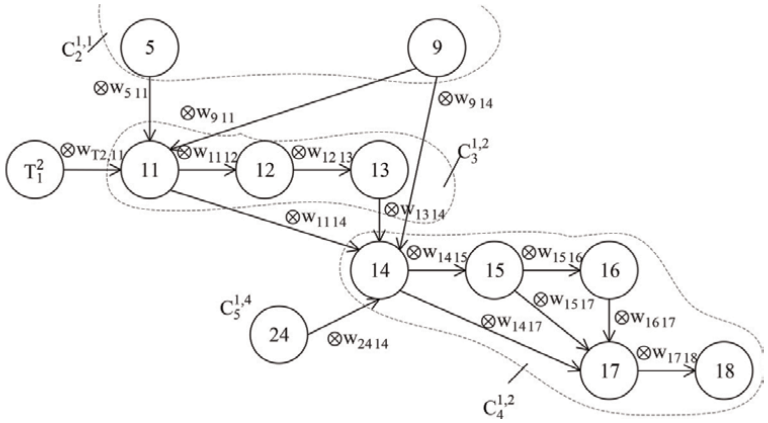
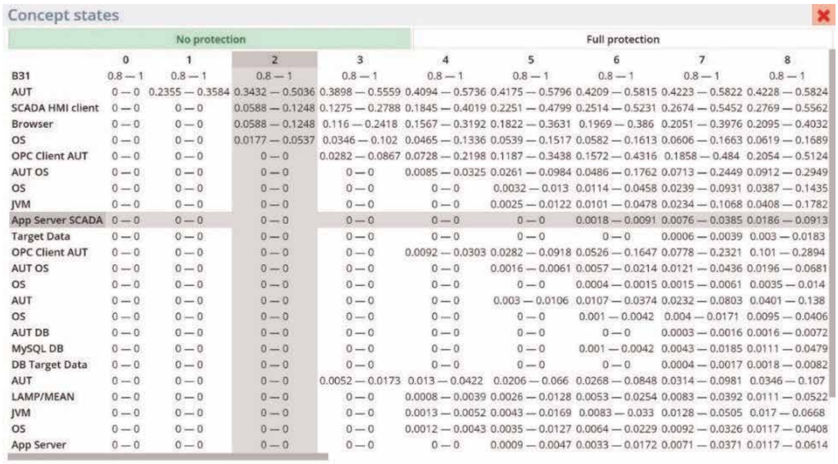


Figure 8.
Decomposition of the concepts $C_3^{1,2}$ and $C_4^{1,2}$ of FGCM for AIS risk assessment


Concept states

	No protection				Full protection				
	0	1	2	3	4	5	6	7	8
B31	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1	0.8 — 1
AUT	0 — 0	0.2355 — 0.3584	0.3432 — 0.5036	0.3898 — 0.5559	0.4094 — 0.5736	0.4175 — 0.5796	0.4209 — 0.5815	0.4223 — 0.5822	0.4228 — 0.5824
SCADA HMI client	0 — 0	0 — 0	0.0588 — 0.1248	0.1275 — 0.2788	0.1845 — 0.4019	0.2251 — 0.4799	0.2514 — 0.5231	0.2674 — 0.5452	0.2769 — 0.5562
Browser	0 — 0	0 — 0	0.0588 — 0.1248	0.116 — 0.2418	0.1567 — 0.3192	0.1822 — 0.3631	0.1969 — 0.386	0.2051 — 0.3976	0.2095 — 0.4032
OS	0 — 0	0 — 0	0.0177 — 0.0537	0.0346 — 0.102	0.0465 — 0.1336	0.0539 — 0.1517	0.0582 — 0.1613	0.0606 — 0.1663	0.0619 — 0.1689
OPC Client AUT	0 — 0	0 — 0	0 — 0	0.0282 — 0.0867	0.0728 — 0.2198	0.1187 — 0.3438	0.1572 — 0.4316	0.1858 — 0.484	0.2054 — 0.5124
AUT OS	0 — 0	0 — 0	0 — 0	0 — 0	0.0085 — 0.0325	0.0261 — 0.0984	0.0486 — 0.1762	0.0713 — 0.2449	0.0912 — 0.2949
OS	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0032 — 0.013	0.0114 — 0.0458	0.0239 — 0.0931	0.0387 — 0.1435
JVM	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0025 — 0.0122	0.0101 — 0.0478	0.0234 — 0.1068	0.0408 — 0.1782
App Server SCADA	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0018 — 0.0091	0.0076 — 0.0385	0.0186 — 0.0913
Target Data	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0006 — 0.0039	0.003 — 0.0183
OPC Client AUT	0 — 0	0 — 0	0 — 0	0 — 0	0.0092 — 0.0303	0.0282 — 0.0918	0.0526 — 0.1647	0.0778 — 0.2321	0.101 — 0.2894
AUT OS	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0016 — 0.0061	0.0057 — 0.0214	0.0121 — 0.0436	0.0196 — 0.0681
OS	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0004 — 0.0015	0.0015 — 0.0061	0.0035 — 0.014
AUT	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.003 — 0.0106	0.0107 — 0.0374	0.0232 — 0.0803	0.0401 — 0.138
OS	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.001 — 0.0042	0.004 — 0.0171	0.0095 — 0.0406
AUT DB	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0003 — 0.0016	0.0016 — 0.0072
MySQL DB	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.001 — 0.0042	0.0043 — 0.0185	0.0111 — 0.0479
DB Target Data	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0004 — 0.0017	0.0018 — 0.0082
AUT	0 — 0	0 — 0	0 — 0	0.0052 — 0.0173	0.013 — 0.0422	0.0206 — 0.066	0.0268 — 0.0848	0.0314 — 0.0981	0.0346 — 0.107
LAMP/MEAN	0 — 0	0 — 0	0 — 0	0 — 0	0.0008 — 0.0039	0.0026 — 0.0128	0.0053 — 0.0254	0.0083 — 0.0392	0.0111 — 0.0522
JVM	0 — 0	0 — 0	0 — 0	0 — 0	0.0013 — 0.0052	0.0043 — 0.0169	0.0083 — 0.033	0.0128 — 0.0505	0.017 — 0.0668
OS	0 — 0	0 — 0	0 — 0	0 — 0	0.0012 — 0.0043	0.0035 — 0.0127	0.0064 — 0.0229	0.0092 — 0.0326	0.0117 — 0.0408
App Server	0 — 0	0 — 0	0 — 0	0 — 0	0 — 0	0.0009 — 0.0047	0.0033 — 0.0172	0.0071 — 0.0371	0.0117 — 0.0614

Figure 9.
FGCM concepts states for risk assessment of Zone 1 (the change in the state of concepts over time and the final states of the target concepts of the FGCM, software window form).

Concept	Concept name	Parent concept
6	Exploitation of the vulnerability of authorization system of the primary OS user	$C_2^{1,1}$
7	Exploitation of the vulnerability of access to operating system memory	
8	Exploitation of the vulnerability of Java virtual machine	
9	Exploitation of the vulnerability of system software of application server for running the SCADA server Web application	
10	The target concept of access to operative TMI data, which can be modified before adding to the database on the nodes of SCADA client-server type	

Table 4.
List of the concepts of the third decomposition level for AIS risks assessment of Zone 1.

Using for calculation the "Cognitive Map Constructor" tool, which is described more detailed in the next section of this chapter, we will estimate the change in the upper and lower boundaries of the state variable X_5 over time $k = 1, 2, 3, \ldots$

Concept	Concept name	Parent concept
19	Exploitation of the vulnerability of authorization system of the main OS user	$C_5^{1,4}$
20	Exploitation of the vulnerability of system software implementing work of Apache Web application server, MySQL DBMS, PHP runtime frameworks to support interactive Web pages	
21	Exploitation of the vulnerability of OS memory access	
22	Exploitation of the vulnerability of Java Virtual Machine Memory Access	
23	Exploitation of the vulnerability of Application Server Software	
24	The target concept of unauthorized launching of the module for access to the database of operative storage of TMI at service stations	
25	Exploitation of the vulnerability of authorization system of the main OS user	$C_6^{1,3}$
26	Exploitation of the vulnerability of access to operating system memory	

Table 5.
List of the concepts of the third decomposition level of Zone 1.

Concept	Concept name	Parent concept
14	Exploitation of the vulnerability of authorization system of the main OS user	$C_4^{1,2}$
15	Exploitation of the vulnerability of OS memory access	
16	Exploitation of the vulnerability of authorization system of the main DBMS user	
17	Exploitation of the vulnerability of DBMS memory access	
18	The target concept of unauthorized modification of TMI operative data TMI stored in the database	
11	Exploitation of the vulnerability of authorization system of the client part of OPC Client UA software	$C_3^{1,2}$
12	Exploitation of the vulnerability of authorization system of the main OS user	
13	Exploitation of the vulnerability of OS memory access	

Table 6.
List of the concepts of the third level of FGCM decomposition of Zone 1.

Linguistic meaning of connection strength	Numeric range
Does not affect	0
Very weak	(0; 0.15]
Weak	(0.15; 0.35]
Average	(0.35; 0.6]
Strong	(0.6; 0.85]
Very strong	(0.85; 1]

Table 7.
Evaluation of the strength of communication between concepts.

(**Tables 9** and **10**). The state of the input concept C_{T_1} is defined here as $\otimes X_{T_1}(k) = [0.8;1]$ for all $= 0, 1, 2, \ldots$; the initial conditions for other state variables $\otimes X_1(0) \div \otimes X_5(0)$ are assumed to be zero, i.e., equal to $[0;0]$.

As a result, the steady-state value of the grey state vector $\otimes X$ for FGCM presented in **Figure 6**, i.e., for the concept $C_1^{1,1}$ decomposition is found as

$$\otimes X = \{[0,8;1],[0,43;0,58],[0,28;0,55],[0,20;0,40],[0,06;0,16],[0,24;0,53]\},$$

and the final value for the target concept state is determined by the grey number $\otimes X_5 \in [0,24;0,53]$.

Consider the state of the target concept C_R (**Figure 2**)—the damage caused by the potential violation of TMI integrity in the AIS—after clarifying all weights by the level of decomposition of the original FGCM. Let us assume that the active

Connection weight	The value of the connection weight	Greyness (scatter of assessment)
$W_{T_1 1}$	[0.6; 0.75]	0.075
W_{12}	[0.5; 0.7]	0.1
W_{13}	[0.5; 0.7]	0.1
W_{14}	[0.15; 0.3]	0.075
W_{25}	[0.55; 0.65]	0.05
W_{32}	[0.35; 0.55]	0.1
W_{35}	[0.55; 0.65]	0.05
W_{42}	[0.3; 0.5]	0.1
W_{43}	[0.15; 0.3]	0.075
W_{45}	[0.2; 0.45]	0.125

Table 8.
The values of communications FGCM weights.

k	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$\overline{\mathbf{X}}_i$								
$\overline{X}_1$	0.36	0.50	0.55	0.57	0.58	0.58	0.58	0.58
$\overline{X}_2$	0	0.125	0.28	0.40	0.48	0.52	0.54	0.55
$\overline{X}_3$	0	0.125	0.24	0.32	0.36	0.38	0.39	0.40
$\overline{X}_4$	0	0.054	0.10	0.13	0.15	0.16	0.16	0.16
$\overline{X}_5$	0	0	0.093	0.23	0.36	0.45	0.50	0.53

Table 9.
Upper boundaries of concept state estimates

k	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$\underline{\mathbf{X}}_i$								
$\underline{X}_1$	0.24	0.34	0.39	0.41	0.43	0.43	0.43	0.43
$\underline{X}_2$	0	0.059	0.13	0.18	0.22	0.25	0.27	0.28
$\underline{X}_3$	0	0.059	0.115	0.16	0.18	0.19	0.20	0.20
$\underline{X}_4$	0	0.018	0.034	0.046	0.052	0.058	0.06	0.06
$\underline{X}_5$	0	0	0.034	0.087	0.14	0.18	0.21	0.24

Table 10.
Lower boundaries of concept state estimates.

threat is the internal threat of violation of the integrity of TMI, the value of which is determined by a grey number $\otimes X^*_{T_1} \in [0,6;0,95]$.

Risk assessment because of violation of the integrity of TMI information is defined as $\otimes X^*_R |_A \in [0.19;0.28]$.

To reduce the potential damage from the violation of TMI integrity, a monitoring system, deployed as a protected container in Zone 5, is used. In **Figure 3**, this information protection tool is designated as a TMI integrity monitoring model—concept IST^5. The protected container ensures the continuous operation of the TMI integrity monitoring system, which implements online and offline analysis of operational data and data collected in the repository (Zone 3).

The concept of TMI integrity monitoring system as a whole has some peculiarities:

- Simulated parameters of the aviation engine operation and TMI can be presented in the form of multidimensional technological time series;

- Monitoring the TMI integrity is based on the analysis of the consistency of the behavior of parameters obtained by using the model of complex technical object, and taken from the on-board aircraft systems;

- The output of the monitoring system is the evaluation of conditional probability of the events of data integrity violation events.

Risk value estimate due to violation of TMI information integrity after applying the tool based on the integrity monitoring model is $\otimes X^*_R |_A \in [0.07;0.15]$.

Due to the significant amount of computation when working with FGCM containing a large number of concepts, it was necessary to develop a software tool to automate cognitive modeling with use of FGCM. The change in the state of concepts over time and the final states of the target concepts of the FGCM, calculated in the developed software tool, are presented in **Figure 9**.

4. Automation of risk analysis and management on the base of cognitive modeling technology

To improve an efficiency of risk analysis and management with use of FGCM, the special software tool "Cognitive Map Constructor" was developed. This software allows us to build and edit FGCM, use them to carry out the security risk analysis, and justify the choice of the necessary countermeasures from the given user-specified set. As a result, a diagram of risk assessment is built under various scenarios of countermeasures' implementation and threats' realization.

Besides supporting the FGCM with the installation of connections weights in the form of the upper and lower boundaries, the software allows us the use of linguistic terms of fuzzy logic, as well as setting the weights in the form of "white" crisp numbers. The software has the interface implemented in HTML using CSS, which allows displaying the FGCM and all the necessary accompanying information by the concepts and connections, and also is able to work on any graphical operating system that has a current Web browser.

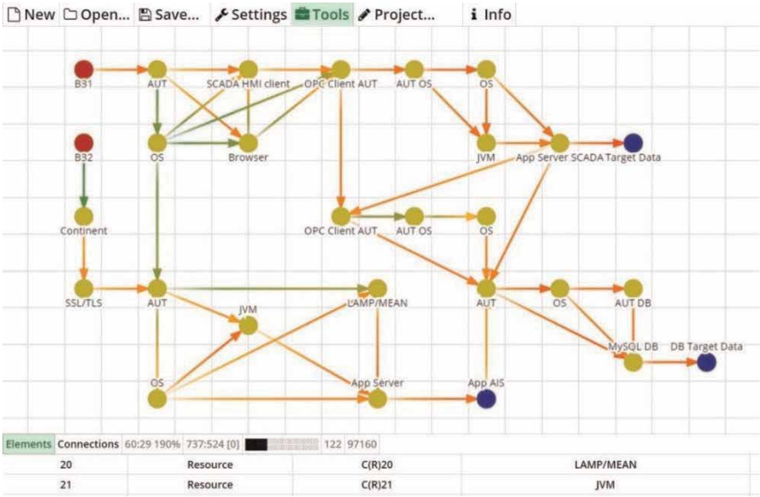
There are five kinds of concepts which are used in FGCM: threats, information assets, intermediate concepts, targets, and countermeasures, which can be marked by different colors for convenience and clarity.

The set of the options depends on the type of the concept, but in most cases its name is specified with description, as well as its current state. In the case when the weights of all connections, pointing to the concept, are assumed to be equal, one can mark the option "Imposed weight" and set the desired value. For countermeasures, it is permissible to indicate which of existing countermeasure it is, that allows realizing situations when one countermeasure acts on several connections at once.

To establish the relationships between the concepts, it is necessary to click on the button "Placement" of the action group "Connections" in the tool window. After that, the connections are located by pressing consecutively on the initial and final element. The located countermeasures and initial states of the concepts can be adjusted and combined, creating the different scenarios that allow us to compare the effectiveness of countermeasures.

Figures 10 and **11** show the FGCM risk estimates built in the "Cognitive Map Constructor."

Thus, the developed software "Cognitive Map Constructor" allows evaluating the effectiveness of the use of the TMI integrity monitoring system in the protection of telemetric information from the effects of external and internal threats.



20	Resource	C(R)20	LAMP/MEAN
21	Resource	C(R)21	JVM

Figure 10.
FGCM for risk assessment of data collection and storage subsystem at the service stations (Zone 1) (software window form).

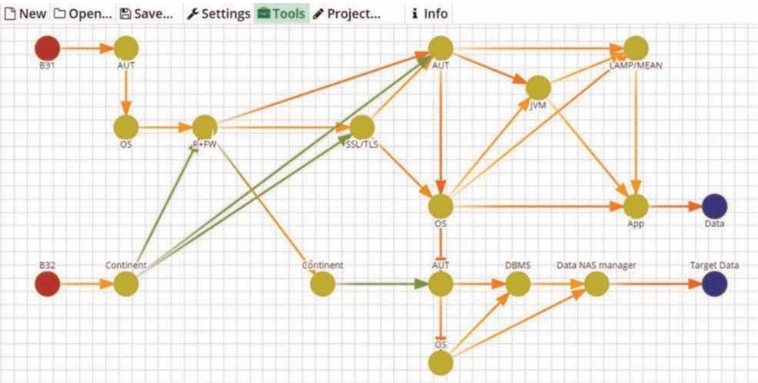


Figure 11.
FGCM for risk assessment in the core of the CIN (Zone 2) and TMI (software window form).

5. Conclusions

A promising way to solve the problem of assessing the cybersecurity risks of industrial automated systems is to model the threats realization scenarios using the tools of topological analysis of the system security and cognitive modeling with the aid of Fuzzy Grey Cognitive Maps.

At the basis of this approach, the construction of original FGCM is proposed to assess the risk of automated control system with the following decomposition of FGCM into the number of cognitive maps of the next level of detail (the same as it is done in IDEF0 Functional Modeling technology). The features of construction of this procedure are discussed in this chapter in relation to the task of ensuring the telemetry information integrity in the industrial automated system for collecting, storing, and processing information on the conditions of on-board aviation systems. It is shown that the use of FGCM allows us to obtain more reliable estimates of security risk factors with account of the possible variations of the available actual data and expert opinions.

To automate the proposed risk assessment procedure in the considered system for collecting, storing, and processing telemetry information with use of FGCM, the software tool “Cognitive Map Constructor” was developed, which can be used for identifying the most dangerous vulnerabilities in the system and evaluating the effectiveness of various measures (countermeasures) realization for telemetric information protection from the impact of external and internal threats.

Acknowledgements

The reported study was funded by RFBR according to the research Project No. 18-00-00238 “Decision support methods and models for innovative project management based on knowledge engineering.”

Author details

Vladimir I. Vasilyev*, Alexey M. Vulfin and Liliya R. Chernyakhovskaya
Ufa State Aviation Technical University, Ufa, Russian Federation

*Address all correspondence to: vasilyev@ugatu.ac.ru

References

[1] WannaCry. Threat Landscape for Industrial Automation Systems in H2 2017 [Internet]. Available from: https://ics-cert.kaspersky.com/tag/wannacry/ [Accessed: 10 May 2019]

[2] Threat Landscape for Industrial Automation Systems. H2 2018 [Internet]. Available from: https://ics-cert.kaspersky.com/reports/2019/03/27 [Accessed: 10 May 2019]

[3] Cyber Security Standards [Internet]. Available from: https://en.wikipedia.org/wiki/Cyber_security_standards [Accessed: 10 May 2019]

[4] Byres E. Using ISA/IEC 62443 Standards to Improve Control System Security. Tofino Security White Paper. Version 1.2. Deutschland GmbH: Tofino Security, a Belden Brand, Belden Inc.; 2014

[5] Salmeron JL. Modelling grey uncertainty with fuzzy grey cognitive maps. Expert Systems with Applications. 2010;**37**(12):7581-7588

[6] Salmeron JL, Papageorgiou EI. Chapter 14: Using Fuzzy Grey Cognitive maps for industrial processes control. In: Fuzzy Cognitive Maps for Applied Sciences and Engineering: From Foundations to Extensions and Learning Algorithms. Intelligent Systems Reference Library. Vol. 54. Berlin/Heidelberg: Springer; 2014

[7] Salmeron JL. Fuzzy grey cognitive maps-based intelligent security system. In: 2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS); August 12–20, 2015; Leicester, UK; 2015. pp. 29-32

[8] Shishkin VM, Savkov SV. The method of interval estimation of risk-analysis system. In: Proceedings of the Second International Conference on Security of Information and Networks (SIN'09); October 6–10, 2009; Famagusta, North Cyprus; 2009. pp. 3-7

[9] Hajek P, Prochazka O. Interval-valued fuzzy cognitive maps with genetic learning for predicting corporate financial distress. Univerzitet u Nišu. 2018;**32**(5):1657-1662

[10] Vasilyev VI, Vulfin AM, Guzairov MB, Kirillova AD. Interval evaluation of information risk with the aid of fuzzy grey cognitive maps. Information Technology. 2018;**24**(10): 657-664

[11] Harmati IA, Koczy LT. On the Convergence of Fuzzy Grey Cognitive Maps. Information Technology, Systems Research, and Computational Physics. Springer Verlag; 2018. pp. 74-84

[12] Zhang JY, Liu ZQ, Zhou S. Quotient FCMs—A decomposition theory for fuzzy cognitive maps. IEEE Transactions on Fuzzy Systems. 2003; **11**(5):593-604

[13] D'Agostini C. Role and meaning of subjective probability: Some comments on common misconceptions. AIP Conference Proceedings. 2001;**568**(1): 23-30. Available from: https://aip.scitation.org/doi/pdf [Accessed: 10 May 2019]

Chapter 3

Application of Chaos-Based Fragile Watermarking to Authenticate Digital Video

Rinaldi Munir and Harlili Harlili

Abstract

Fragile watermarking algorithm is a technique used to authenticate of digital data multimedia such as video. A watermarking algorithm consists of two processes: embedding and extraction of watermark. In this paper, a secure video fragile watermarking algorithm in spatial domain based on chaos is proposed. The watermark is a binary image which has the same size with frame size of the video. Before embedding, in order to increase security, the watermark is encrypted using XOR operation with a random image. The random image is generated by using Cross-Coupled Chaotic random Bit Generator (CCCBG). The encrypted watermark is embedded to each frame. In the extraction process, the encrypted watermark is extracted from the watermarked video, decrypted it, and then compared to the original watermark. This algorithm has capability to localize the area being tampered in the video frames. We have performed some typical attacks to the watermarked video and then authenticated it. Based on the experiment results, the algorithm can detect and localize the modified region of video frames very well. Sensitivity to the slightest change on initial conditions of the chaos map provided security of the algorithm.

Keywords: fragile watermarking, authentication, digital video, chaos

1. Introduction

Nowadays digital video is widely used to present information. This is because a video is richer in information compared to a single image, and generally a video has more pictures and sound. However, digital data such as video have advantages and disadvantages. Digital video could be edited, manipulated, or altered easily by using a video editor or other tools. For example, someone could change contrast of the video, resize, remove or add some frames, or add a new object to the video. Unfortunately, once a digital video is manipulated, its integrity is questionable. In some cases, we need to know authenticity of the video. For example, a court need to decide if a video as evidence is genuine or has been manipulated. If the video has been manipulated, how to prove it?

The integrity problem of digital video could be solved by using a fragile watermarking technique. In the fragile watermarking technique, we could embed one or more watermarks into frames of video. Once the watermarked video is manipulated, altered, or modified, the watermark inside will be fragile or damage. The damaged watermark is indication that the video has been manipulated. Therefore, fragile watermarking can be used to prove authentication of a video.

IntechOpen

A watermarking algorithm consists of two processes: embedding and extraction of watermark. Watermark is an information that refers to the video owner. Usually the watermark is a binary image such as logo, random bits, or other information. The watermark is inserted into a host video, frame by frame, become to a watermarked video without affecting its perceptual (and audio) quality. Through an inverse process, the embedded watermark can be extracted again from the video. When the extracted watermark is compared to the original watermark, we could conclude if the video has been altered. The damage extracted watermark is indication that the video has been altered.

Digital watermarking schemes can operate on spatial domain or frequency domain. Assume the watermark is represented as string of bits. Watermarking schemes that operate in spatial domain embed watermark bits into pixel values of the video frame directly [1, 2]. Otherwise, on watermarking schemes that operate in transform domain, a host video has to be transformed first into a transform domain by using a specific transformation (DCT, DWT, DFT, etc.) [3]. Next embedding bits of watermark is performed by modifying the transform coefficients [2, 4].

Generally watermarking in transform domain is more robust than spatial domain through non-malicious attacks such as cropping, compression, scaling, rotation, etc. Therefore, it is used to the problems of copyright protection, proving ownership, illegal copying, and transaction tracking of video. Otherwise, watermarking in spatial domain is suitable to solve the problem of tamper detection of video content. The watermark in the video must be fragile when the video is manipulated. Robustness is not important for fragile watermarking.

The watermark could be originated from internal or external. The internal watermark means that the watermark is derived from host video directly, and then it is embedded into frames of the video. The second is external watermark which means that the watermark is an input from the user, and usually the watermark is a meaningful binary image such as logo or other image.

Much research on fragile watermarking for digital video has been done by many scientists. This means that the fragile watermarking is interesting research topics. Some related research is from Elgamal et al. [1], Zhi-yu et al. [4], and Rupali et al. [5]. Elgamal et al. [1] proposed a fragile video watermarking algorithm on transform. The original video is transformed from RGB model to YCbCr model and then Cr-component is partitioned into non-overlapping blocks of pixel. The watermark is a binary image from the video owner. Bits of the binary image are embedded for each block separately.

Like Elgamal et al. [1], Zhi-yu et al. [4] also proposed a fragile watermarking algorithm on transform domain. The original video is transformed from RGB model to YST model. The T-component is divided 4 × 4 blocks and then each block is transformed to frequency domain using DCT. The watermark is generated from the quantized DCT coefficient and then it is embedded into the last non-zero DCT coefficient.

Rupali et al. [5] also proposed a fragile video watermarking algorithm on transform domain. The original video is transformed from spatial domain to frequency domain using DCT. Rupali et al. [5] used two watermarks to embed. Both of these watermarks are internal, that means from the original video itself. The first watermark is used to detect tampering and the second watermark is used to localize tampered area.

All of the watermarking algorithms above operated on transform domain. A digital video contains more frames, generally hundreds to thousands frames. Transformation of each frame from the spatial domain to the transform domain consumes a lot of computing time. We need a simple fragile watermarking for the digital video but still meet the security aspect. A simple fragile watermarking on spatial domain is by using LSB (least significant bit) modification method. This method is fast and it can detect tampering on video until pixel level. To fulfill the security aspect, we use the watermarking key(s) in embedding and extraction process, so that embedding and extraction of watermark are performed by an authorized party who has the secret key(s).

The watermark itself is confidential, only the video owner knows about it. Therefore, the watermark needs to be encrypted using the secret key(s). The secret key(s) also serve to prevent the watermark from being extracted and used in the reassembling of videos by an authorized party, thus avoiding counterfeiting of the videos.

The chaos system can be used to get a secure fragile watermarking scheme. In recent years, chaos theory has attracted the attention of scientists, especially in the information security field. Chaos has been used to increase security [6]. The reason is the chaotic systems that have sensitivity on initial conditions. It means if we perform a little bit change to the initial conditions of the chaos system, after some iterations, the system will result values that differ significantly. In the field of cryptography and watermarking, generally a chaos map is used to generate pseudo-random numbers [7, 8].

Munir et al. [9] used a chaos map, i.e. Arnold Cat's map, to encrypt the watermark before embedding it into video frames. The original watermark is encrypted by XOR-ing it with a random image. The random image is generated from the replicated watermark by using an Arnold Cat Map. The encrypted watermark is embedded into each frame of the video using LSB modification method.

Unfortunately, the random image generated using Arnold Cat's map still shows patterns of the replicated watermark, so it is not completely random. The embedding algorithm is also redundant, because the random image is XOR-ed with the replicated watermark. Therefore, the watermarking algorithm has redundancy.

In this paper, we modified the previous algorithm by using another chaos map so that the random image is generated from the chaos map directly. We used a random bit generator based on two Skew Tent Maps. The generator is abbreviated as CCCBG (Cross-Coupled Chaotic random Bit Generator) [10].

The paper is organized into six sections. The first section is this introduction. The second section will review some supported theories. The algorithm to embed and extract the watermark will be explained in the third section. The fourth and fifth section will present the experiment results and discussion. Finally, the last section will resume the conclusion and future work.

2. Chaos map

One of the popular chaotic maps is a Logistic Map, described by

$$x_{i+1} = \mu x_i \left(1 - x_i\right) \tag{1}$$

where μ is a parameter of map and $0 < \mu \leq 4$. According to [7], the map is in chaotic state when $3.57 < \mu \leq 4$. In this state, the resulting values appear random. Because of its random behavior, a logistic map can be used as a pseudo-random generator. Hence, initial value of the chaos map, x_0, and constant μ serve as secret keys. When we iterate Eq. (1) from an initial value (x_0), we get a random sequences between 0 and 1. The random values generated from the chaos Logistic Map are sensitive to small changes of the initial values. By changing x_0, the random values generated different significantly from the previous chaotic values with initial value x_0.

Another chaos map is Tent Map. It iterates a point x_0 and gives a sequence x_i in [0, 1]:

$$x_{i+1} = f_\mu(x_i) = \begin{cases} \mu x_i & , x_i < \frac{1}{2} \\ \mu(1 - x_i) & , x_i \geq \frac{1}{2} \end{cases} \tag{2}$$

where μ is a positive real constant. Varian of Tent Map is Skew Tent Map [8] which is defined as:

$$x_{i+1} = f_\mu(x_i) = \begin{cases} \frac{x_i}{\mu} & , 0 \leq x_i < \mu \\ \frac{1 - x_i}{1 - \mu} & , \mu < x \leq 1 \end{cases} \tag{3}$$

where μ is system parameter and x_0 is initial condition of map. When a Skew Tent Map is iterated from x_0 value, it produces a sequence in the interval [0, 1] and distributed uniformly.

Narendra et al. in [10] proposed a random bit generator based on two Skew Tent Maps. The generator is abbreviated as CCCBG (Cross-Coupled Chaotic random Bit Generator). In the CCCBG, random bit stream is generated by comparing outputs of the couple maps. If $f_\mu(x_i)$ and $g_\mu(y_i)$ are two Skew Tent Maps and are given as:

$$x_{i+1} = f_\mu(x_i) \tag{4}$$

$$y_{i+1} = g_\mu(y_i) \tag{5}$$

where μ is system parameter and is same for both maps. The CCCBG generated a sequence of random bits by comparing the outputs of the maps in the following way:

$$h(x_{i+1}, y_{i+1}) = \begin{cases} 0 & , x_{i+1} < y_{i+1} \\ 1 & , x_{i+1} \geq y_{i+1} \end{cases} \tag{6}$$

Based on several tests performed by Narendra et al., the CCCBG successfully passes all the randomness tests [10], therefore the random binary sequences can be used for encryption.

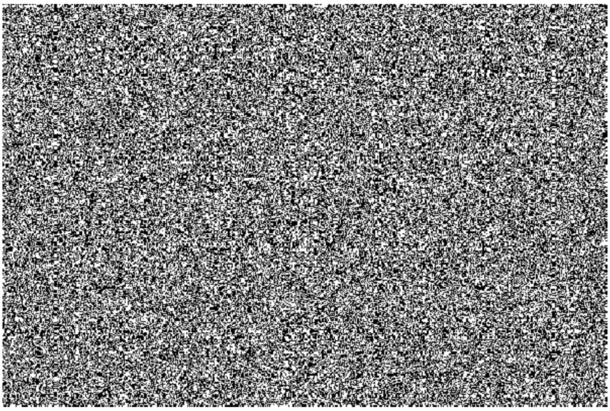


Figure 1.
A random image generated by using the CCCBG.

In digital watermarking, the random bits play important role to increase security. We will use CCCBG Map to generate random bits. The random bits will be-XOR-ed to the original watermark to yield the encrypted watermark. For example, by iterating (6) 240,000 times (i.e., 480 × 854) and using parameter $\mu = 0.48999$, initial conditions $x_0 = 0.500684$, and $y_0 = 0.538167586$, we get a sequence of random bits figured as a binary bit image (**Figure 1**).

3. Proposed algorithm

This section will explain the proposed fragile video watermarking on spatial domain based on the chaotic map. The watermark is a binary image. To detect manipulation in the video frames until pixel level, the watermark must have the same size as the video frame size. Therefore, if watermark size is less than video frame size, the watermark need to be replicated by duplicating it a number of times in order to produce a new watermark that has the same size with the host video frame size. **Figure 2** shows example of replication. The original watermark "ASEAN logo" has a size of 200 × 194 pixels, whereas the video frames have a size of 480 × 854 pixels. This original watermark must be duplicated a number of times so that produce a replicated watermark that has size 480 × 854.

Next, to increase security, before embedding, the replicated watermark is encrypted by XOR-ing it with a random image. A random image is generated by iterating CCCBG a number of mn times where m and n are frame sizes (**Figure 1**). The replicated watermark is encrypted with the random image by using XOR operation to produce an encrypted watermark (**Figure 3**). Next, we embed the encrypted watermark into the host video.

We design a simple, but secure, fragile video watermarking based on chaos. The fragile video watermarking algorithm consists of two processes: embedding algorithm and extraction algorithm, each will be described below.

3.1 Watermark embedding algorithm

There are two scenarios for embedding the watermark into a digital video. The first scenario is embedding each frame of the video with the same watermark. The second scenario is embedding each frame of the video with the different watermark. The second scenario is not practical because the video owner have to provide the watermark of as many frames. Actually, the watermarks can be generated from the video itself (i.e., internal watermarks), we generate the watermark for each video frame that

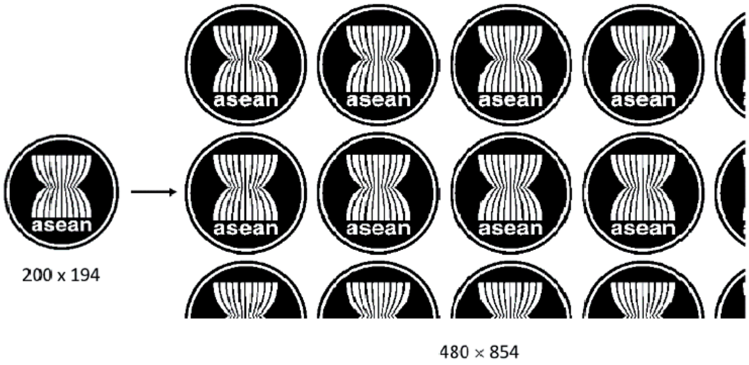


Figure 2.
Left: the original watermark; right: the replicated watermark.

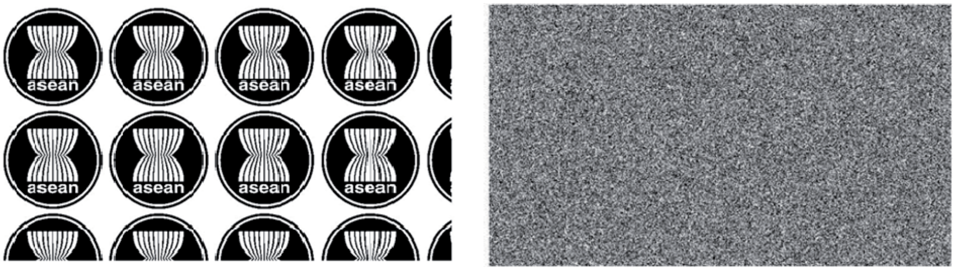


Figure 3.
Left: the replicated watermark. Right: the encrypted watermark.

depend on the frame content itself. However, the resulting watermark is meaningless and cannot be perceived visually. We want the watermark to be meaningful and can be perceived visually. Therefore, we choose the watermark is a meaningful binary image and, for practical reason, the same watermark is embedded to each video frame.

Now, we can describe the watermark embedding algorithm into the digital video in more detail as follows:

Input: a host video file (v), a watermark file (w), and CCCBG's parameter and initial conditions (μ, x_0, y_0).

Output: a watermarked video (v').

Step 1: Read the frames of video v, the watermark w, and CCCBG's parameter and initial conditions (μ, x_0, and y_0). If the video has an audio, then separate the audio.

Step 2: If size(w) < size(video frame of v), copy the single watermark to produce a replicated watermark w' which has the same size with the host video frames.

Step 3: Iterate CCCBG mn times to produce a random image r.

Step 4: Encrypt w' by XOR-ing it with r as follows:

$$w'' = w' \oplus r \tag{7}$$

Step 5: Embed the encrypted watermark, w", into each frame of the video by manipulating the least significant bit (LSB) of pixels. If the frame has R, G, and B component, then it performs embedding to each component.

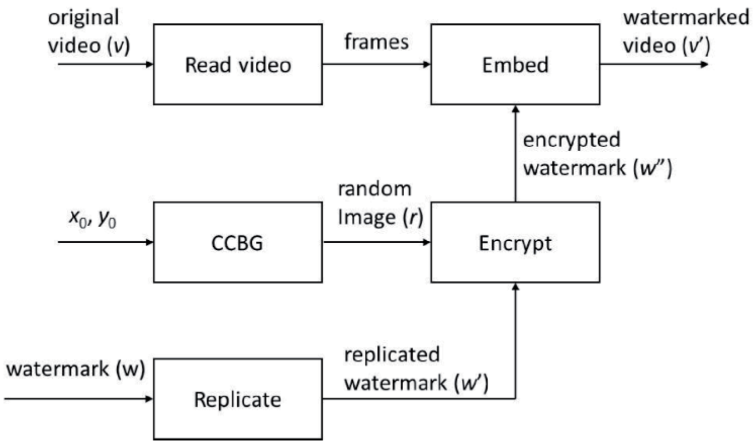


Figure 4.
Watermark embedding algorithm.

Step 6: If the original video has audio, merge it to the watermarked frames to produce a watermarked video.

Figure 4 shows stages in the watermark embedding algorithm. The watermark is embedded into each frames of video. If the video has audio, the audio has not been changed. After embedding of the watermark, audio is merged back into the video.

3.2 Watermark extraction algorithm

Watermark extraction from the video is performed to prove video authentication. Only the video owner can do it, because the owner has the original watermark. The original watermark is required to compare it with the extracted watermark. If the extracted watermark is the same as the original watermark (can be observed visually), then we decide that the video is authentic, otherwise the video has been altered, tampered, or manipulated. The original watermark is also required to localize the tampered region in a frame. Watermark extraction can be performed on all video frames or only on certain frames.

Now, we can describe the watermark extraction algorithm from the digital video in more detail as follows:

Input: a watermarked video file (v'), an original watermark file (w), and CCCBG's parameter and initial conditions (μ, x_0, and y_0).

Output: an extracted watermark, location of tampered frame (if any).

Step 1: Read the frames of video v', watermark w, and CCCBG's parameter and initial conditions (μ, x_0 and y_0).

Step 2: If size(w) < size(video frame of v), copy the single watermark to produce a replicated watermark w' which has the same size with the host video frames.

Step 3: Iterating CCCBG mn times to produce a random image r.

Step 4: For each frame, extract all of the least significant bit (LSB) of pixels. This step yields an extracted watermark w".

Step 4: Decrypt the watermark w" by XOR-ing it with r as:

$$w''' = w'' \oplus r. \tag{8}$$

Step 5: Compare w"' with w'. If w"' = w', we conclude that the integrity of video is authenticated. If not, go to step 6 and 7.

Step 6: To localize tampered region, subtract w' to w"'. If a pixel is not changed, the subtraction yields 0, else the subtraction yields 1.

Step 7: Identify pixels in the watermarked framed in position where the subtraction above yields 1. Those are pixels that have been manipulated.

Figure 5(a) shows stages in the watermark extraction algorithm. The watermark is extracted from each frame of video and then compares the extracted watermark to the original watermark. The decision is binary (1 or 0), 0 means the watermarked video has not changed (authentic), 1 means the watermarked video has been manipulated, altered, or tampered. **Figure 5(b)** shows how to localize tampered region in a frame.

4. Experiment and results

We have implemented the proposed algorithm to be a computer program. Next we test the algorithm on a sample video to determine authentication of the video. The

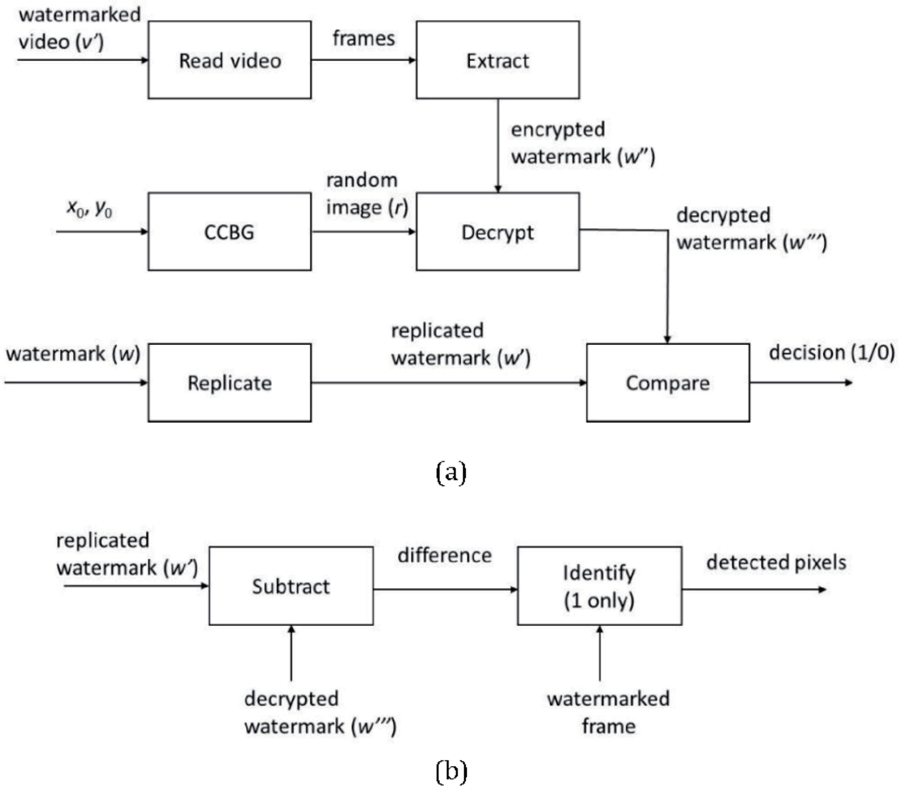


Figure 5.
(a) Watermark extraction algorithm and (b) localize tampered region.

sample video was a video clip which has 394 frames and long 16 seconds, each frame has a size 480 × 854 (**Figure 6a**). This video contains audio inside. **Figure 6b** and **c** show two frames of the video (frame 1 and frame 175).

The watermark to be embedded into the host video is “ASEAN logo” as shown in **Figure 3** (after replicated). In these experiments below, we used parameters of CCCBG as follows: $\mu = 0.48999$, initial conditions $x_0 = 0.5006841$, and $y_0 = 0.538167586$. The two initial conditions serve as the secret keys. These keys were used in both watermark embedding algorithm and extraction algorithm.

This algorithm can only be used to test the authentication of digital videos that have been spatially manipulated. Spatial manipulation such as changing the contrast of the image, adding noise, changing the size of the frame, copy and paste an object into the frame, and others. The algorithm cannot temporarily detect video manipulation, for example, by removing one or more frames.

In the section below, we divide experiments into two cases: (i) no attack case and (ii) tamper detection test, each will be described below.

4.1 No-attack case

If the watermarked video is not manipulated, we categorized it as no attack case. To prove the authentication of the video, we extracted all watermarks from each video frame. All watermarks should be the same as the original watermark. **Figure 7** shows the extracted watermarks but only from frame number 1 and frame number 283.

Visually, there are no damages in the extracted watermarks. When we compare the extracted watermark to the original watermark by subtracting them, we get

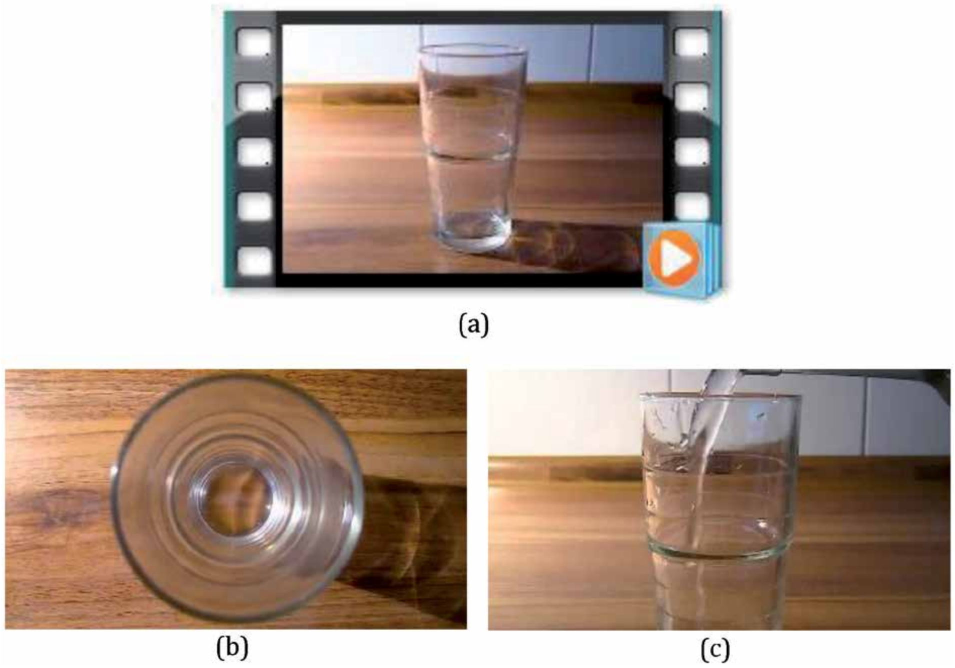


Figure 6.
(a) A host video to be watermarked, (b) frame 1, and (c) frame 175.

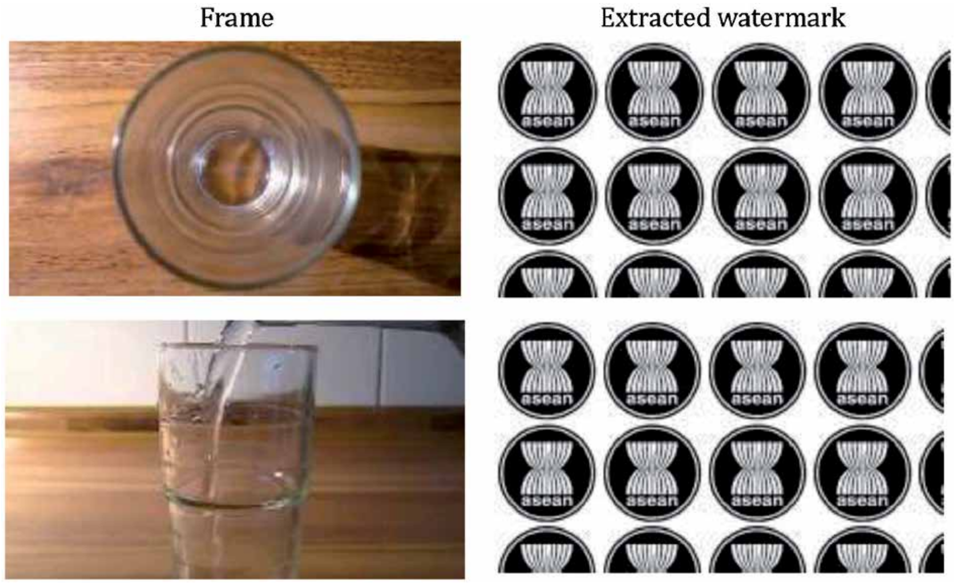


Figure 7.
The watermarked frames and the extracted watermarks.

the difference is a black image (all of pixels are 0). Therefore, we conclude no tampering performed to the watermarked video. In this algorithm, parameter and initial condition of CCCBG behave as secret keys. Embedding and extraction of the watermark could be done by the authorized party only. If the receiver did not have the same keys, then the extracted watermark is not the same as the original watermark.

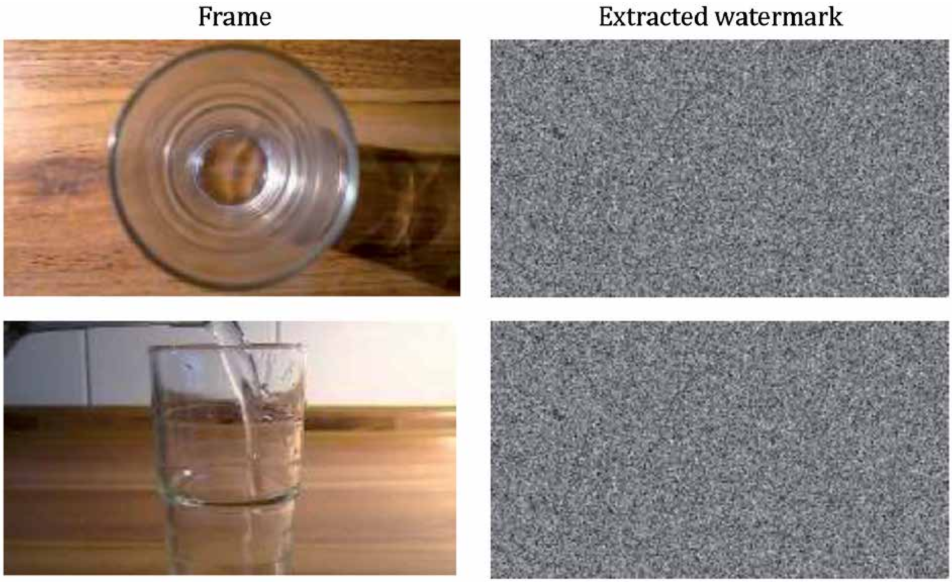


Figure 8.
The watermarked frames and the extracted (wrong) watermarks.

4.2 Sensitivity to initial condition

As mentioned before, the chaos system has a sensitivity to the slightest change on initial conditions. This characteristic provides the security aspect of watermarking. For example, the receiver used μ = 0.4900 (before was 0.48999), initial conditions x_0 = 0.5006840 (before was 0.5006841), and y_0 = 0.538167585 (before was 0.5381675865) to extract the watermark. **Figure 8** shows the extracted watermarks from two frames. The extracted watermarks are wrong! Compared to the original watermark, this extracted watermarks look like the random images. This happens because CCBG produces random bits that are very different from previous bits.

4.3 Tamper detection test

In most cases, a digital video is often edited or manipulated using the video editor. If a video has been manipulated, the video is no longer original. Main goal of fragile watermarking is to determine if the video has been manipulated or not. If the video has been manipulated, the algorithm should able to locate where the alteration made on the video frames. In these experiments, we performed some typical attacks to the watermarked video. The attacks are (1) adding a text to the watermarked video, (2) copy-paste attack, (3) adding some noises, (4) modifying video contrast, and (5) cropping the frames. The following are the attacks.

5. Detection test against text addition

We attack the watermarked video by writing a text “GLASS and WATER” at the left top of the frames (**Figure 9a**). To prove the authentication of the video, we extracted the watermarks from the video frames to get the extracted watermarks. The extracted watermarks contain the text (**Figure 9b**). Therefore, we conclude that the watermarked video has been manipulated. **Figure 9c** shows detection of

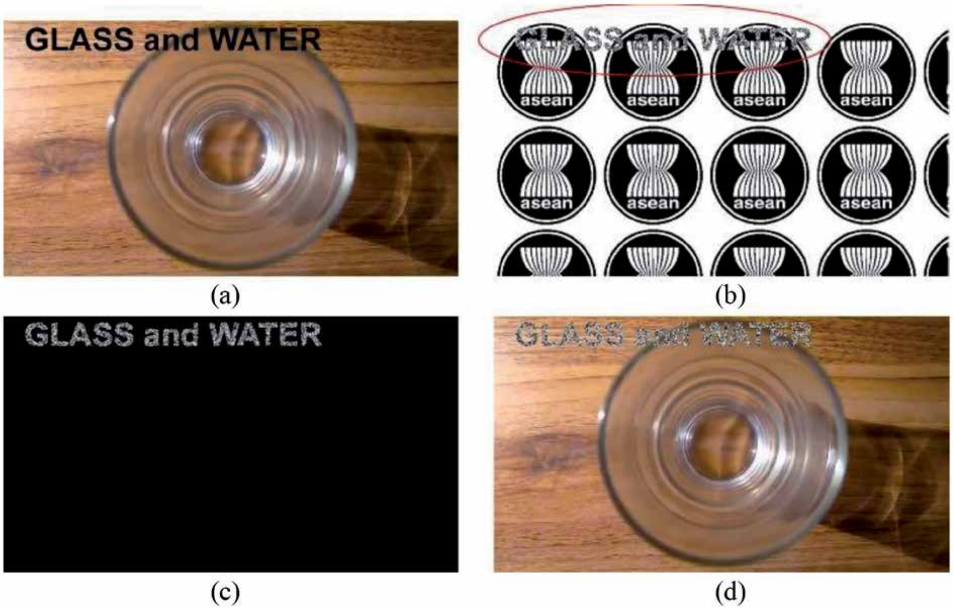


Figure 9.
(a) Watermarked frame after adding a text; (b) extracted watermark; (c) and (d) detected tampering region.

pixels that have been manipulated by adding a text "GLASS and WATER." **Figure 9d** shows the tampered pixels in the correspondence frame.

6. Detection test against copy-paste attack

In the second attack, we copied an object "coca-cola bottle" and then pasted it into the watermarked video (**Figure 10a**). When we extracted the watermarks from the

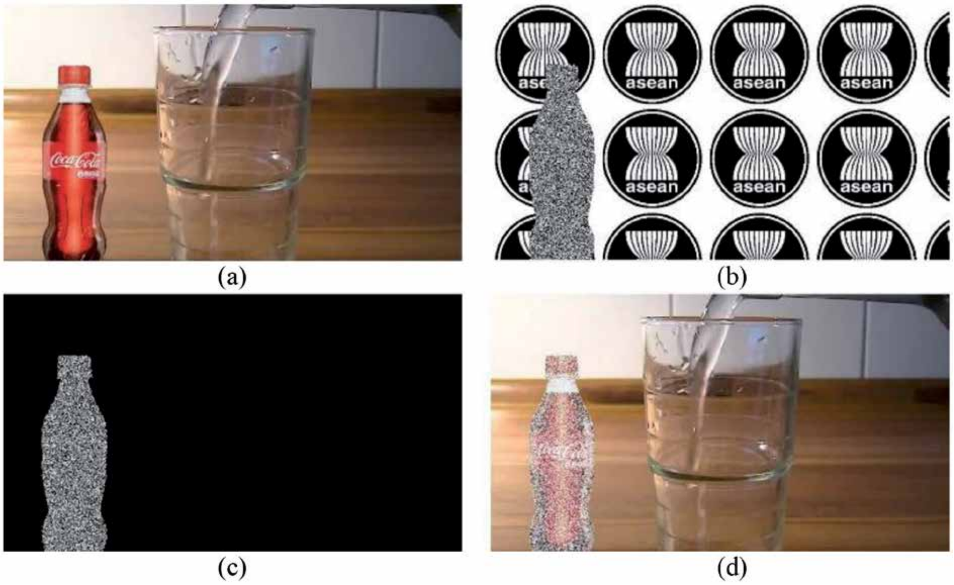


Figure 10.
(a) Watermarked frame after copy-paste attack; (b) extracted watermark; (c) and (d) detected tampering region.

video, we got an extracted watermark as shown in **Figure 10b**. The extracted watermark contains a silhouette of strange object inside. Localize the tampered region and we can detect copy-paste object in the video frames as shown in **Figure 10c** and **d**.

7. Detection test against adding some noises

There are some kinds of noise such as Gaussian noise, salt and pepper noise, Poisson noise, etc. In the third attack, we added “salt and pepper” noise with density 0.1 into the watermarked video (**Figure 11a**). When we extracted the watermarks from the video, we got the watermarks also contained noise. The noisy watermarks indicated that the video has been altered (**Figure 11b**). The tampering region is entire of frame (**Figure 11c** and **d**).

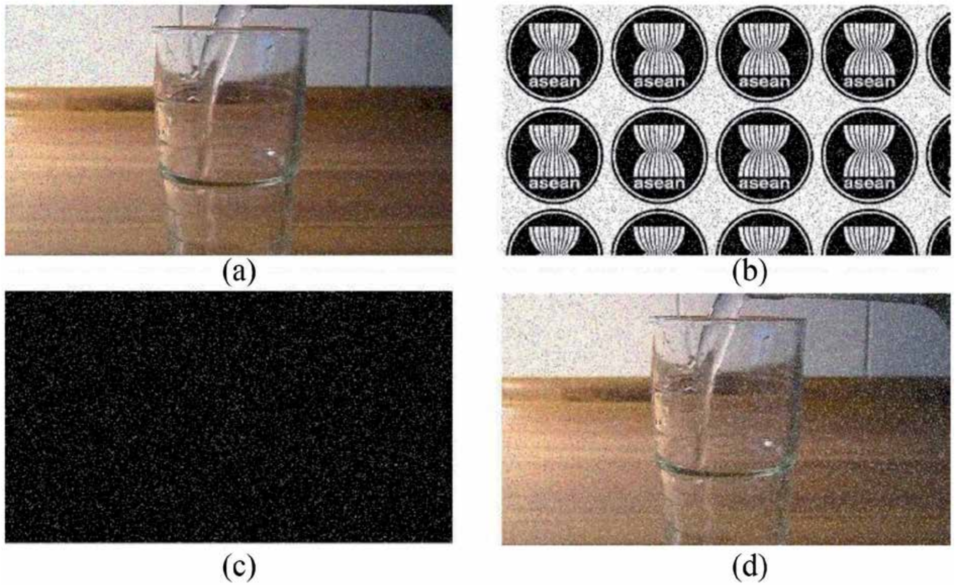


Figure 11.
(a) The watermarked frames after adding noise “salt and pepper”; (b) the extracted watermark; (c) and (d) detected tampering region.

8. Detection test against contrast change

A digital video can be changed so that the contrast becomes brighter or darker. In this attack, we manipulated the watermarked video by changing the contrast so that make it brighter. After that, we extracted the watermarks from the video (**Figure 12**, top right). We can see that the extracted watermarks are damaged and cannot be recognized anymore. Localization of tampered region shows that whole of image has been manipulated (**Figure 12**, bottom right).

9. Detection test against cropping

One of the geometrical attacks is cropping. In this attack, we manipulated the watermarked video by cropping the video frames. The cropping can be performed horizontally or vertically. In this experiment, we cropped a left side of the video. To extract the

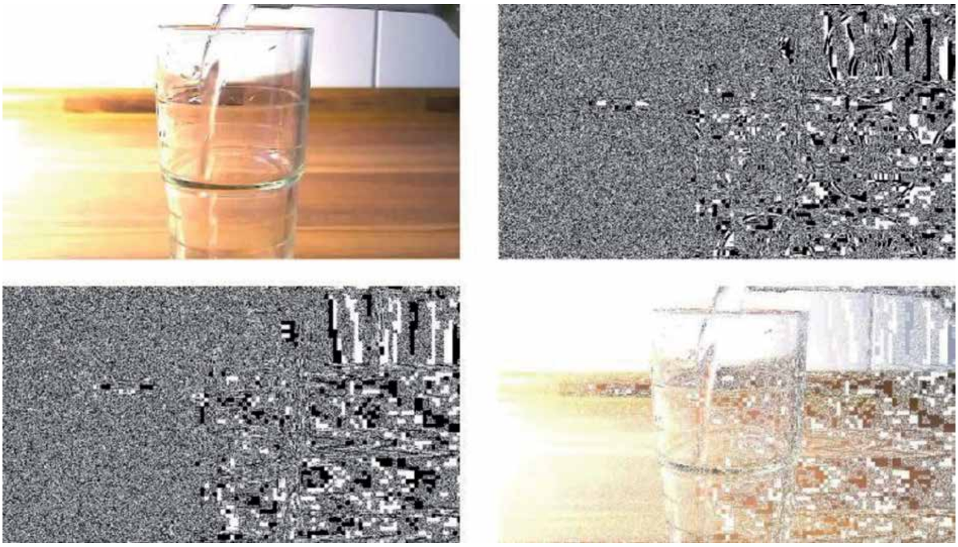


Figure 12.
Top left: the watermarked frame after changing the brightness; top right: the extracted watermark. Bottom right: the tampered region.

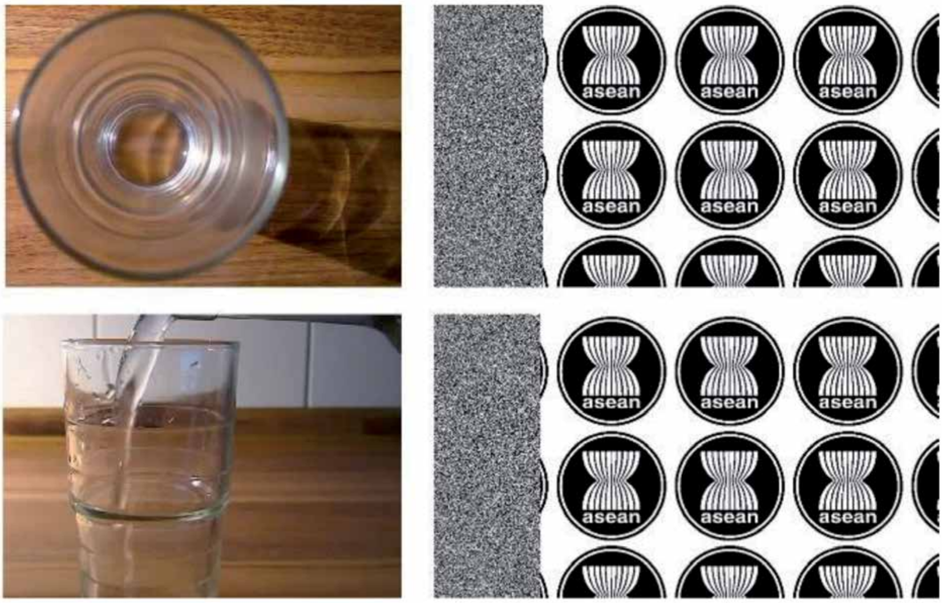


Figure 13.
The watermarked frames after cropping and the extracted watermarks.

watermarks, we returned the frame size into original size first by adding white or black pixels (in this experiment we added white pixels). We found the extracted watermarks contained the black region that indicated the cropped region in the frames (**Figure 13**).

10. Discussion

The proposed chaos-based fragile watermarking algorithm is simple but secure; it can be used to authenticate the digital video. Some experiments have been done to test performance of the algorithm. If there is no manipulation done to the watermarked video (no attack case), then the extracted watermark is same exactly to the original watermark. Therefore, we conclude that the video is still original, has not been changed or manipulated.

Common manipulations of video have been done to test authentication and localize altering in the watermarked video. These manipulations are adding a text label into video frames, inserting a new object into the video, changing contrast, and cropping some pixels. In the case of adding text and inserting an object into the video frames, we got the extracted watermarks that contain silhouette of the object or text. The silhouettes can be seen visually. When we compared to the original watermark, the extracted watermark is not the same. Therefore we conclude that the video has been manipulated. By subtracting the original watermark from the extracted watermark and adjusting the results on the watermarked video, we can find the video frame portion that has been changed.

Common manipulation of video is changing the contrast or brightness of the video. By changing the contrast or brightness of the video, it means changing all pixel values in the video frame. When the watermarks are extracted from the video, we found that extracted watermarks also change entirely. The extracted watermarks are totally damaged; therefore we can conclude that the video has been manipulated.

When a block area in the watermarked video frame is cropped, the extracted watermark is also cropped in the correspondence block. The extracted watermark has a black region in the cropped area of the correspondence frame.

This proposed algorithm has some weakness. It cannot detect manipulation of the watermarked video if one or more fames are removed. However, if some video frames are inserted to the watermarked video, the algorithm can still detect this manipulation, because the new frames do not contain the embedded watermarks.

Other weakness is LSB modification method itself. Bits of the watermark are only embedded to one least significant bit of pixel values. If manipulation of the watermarked video is performed on other than the least significant bit, the algorithm cannot detect it. However, this manipulation is considered uncommon so it can be ignored.

11. Conclusion and future works

We have proposed a fragile video watermarking based on the chaotic map. In order to increase security, the watermark is encrypted using XOR operation with a random image. The random image is generated by using Cross-Coupled Chaotic random Bit Generator (CCCBG). The encrypted watermark is embedded to every RGB component of each frame. In the extraction process, the encrypted watermark is extracted from the watermarked video and compared to the original watermark.

Some experiments have been done to test capability of the algorithm to detect tampering to the watermarked video. We have tried some common attacks to the watermarked video. The experiment results showed that the algorithm could detect tampering on the watermarked video. This algorithm has also capability to localize the area being tampered in the video frames.

The algorithm can only be used to the uncompressed videos. It can be developed for the compressed video format such as MPEG-4. Embedding of watermarks is performed in encoding and decoding must be operated in transform domain. Some transform methods such as Fourier Transform, DCT, or wavelet transform can be used.

This algorithm can also be developed so that it can detect the manipulation that removes one or more frames from the watermarked video. This can be done by using the internal watermarks that depend on the entire video content. Therefore, if some video frames are removed, the internal watermarks also change.

Author details

Rinaldi Munir* and Harlili Harlili
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

*Address all correspondence to: rinaldi@informatika.org

References

[1] Elgamal AF, Mosa NA, ElSaid WK. A fragile video watermarking algorithm for content authentication based on block mean and modulation factor. International Journal of Computer Applications. 2013;**80**(4):0975-8887

[2] Jayamalar T, Radha V. Survey on digital watermarking techniques and attacks watermark. International Journal of Engineering, Science and Technology. 2010;**2**(12):6963-6937

[3] Maryam A, Mansoor R, Hamidreza A. A novel robust scaling image watermarking scheme based on Gaussian mixture model. Expert Systems with Applications. 2015;**42**(4):1960-1971. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0957417414006381

[4] Zhi-Yu H, Xiang-Hong T. Integrity authentication scheme of color video based on the fragile watermarking. In: Proceedings of 2011 International Conference on Electronics, Communications and Control (ICECC). 2011. Available from: https://ieeexplore.ieee.org/document/6067709

[5] Rupali DP, Shilpa M. Fragile video watermarking for tampering detection and localization. Proceedings of 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI). 2015

[6] Dawei Z, Guanrong C, Wenbo L. A chaos-based robust wavelet-domain watermarking algorithm. Chaos, Solitons & Fractals. 2004;**22**:47-54

[7] Bose R, Banerjee A. Implementing symmetric cryptography using chaos function. In: Proceeding 7th International Conference on Advanced Computing and Communication (ADCOM). Indian Institute of Technology; 20 Decembe 1999. pp. 318-321

[8] Stojanovski T, Pihl J, Kocarev L. Chaos-based random number generators - part II: Practical realization. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications. 2001;**48**(3):382-385

[9] Munir R, Harlili H. A secure fragile video watermarking algorithm for content authentication based on Arnold Cat's map. Proceedings of the 4th International Conference on Information Technology (InCIT2019). Bangkok, Thailand; 24-25 October 2019

[10] Narendra KP, Vinod P, Krishan K. A random bit generator using chaotic maps. International Journal of Network Security. 2010;**10**(1):32-38

Section 2

Digital Forensics - Legal Aspects

Chapter 4

Legal Framework on Child Pornography: A Perspective

Deepa Salian and Sofia Khatun

Abstract

Child pornography is a crime which is practiced in every nation of the world. It has also emerged has the fasted growing online business nearly $3 billion annual revenue is been extracted from this sources. The website that tend to distribute commercial child pornography initiates thousands of images and videos and it is extremely difficult to eradicate the images once it is been upload on the internet virtually, so children are lifelong victims of this crime. In the most of the countries Pre- adolescent are involved in child pornography they are subjected to physical as well as sexual violence. It has to be noted that there are two ways in which child pornography is been distributed, i.e., either for profit or for noncommercial child pornography offered through free traded among the offenders. Child pornography is visual delineate act of sexual explicit which involves intercourse, bestiality, masturbation etc. child pornography is an international felony which need to eradicated completely for this every state needs to adopt stringent measure to curb this crime.

Keywords: adolescent, sexual violence, government, sexual maltreatment, communities, entertainment

1. Introduction

The word pornography originates from the Greek pornographic actually significance expounding on whores. One of the ordinarily acknowledged meanings of sex entertainment in current occasions characterizes it as explicitly express material (verbal or pictorial) for example essentially intended to deliver sexual excitement in watchers. At the point when esteem decisions are joined to this definition, sex entertainment is seen as explicitly express material intended to deliver sexual excitement in buyers that is awful with a particular goal in mind.

Child pornography is distributing and transmitting revolting material of children in electronic structure. Lately child pornography has expanded because of the simple access of the web and effectively accessible recordings on the web. Child pornography is the most offensive wrongdoing which happens and has prompted different violations, for example, sex tourism, sexual maltreatment of the youngster, and so on.

Child pornography laws give seriously punishments to makers and wholesalers in practically all Western social orders, ordinarily including imprisonment, with shorter length of sentences for non-business appropriation relying upon the degree and substance of the material circulated.

INTERPOL has referred to Germany as one of the significant makers of child pornography, with the Netherlands and the United Kingdom as the significant dissemination communities. The United States is probably the biggest market of interest for child pornography, however more premiums have moved to Southeast Asia as of late. The improvement of child pornography is fuelled by fundamentally two factors, the beginning, and accessibility of home motion pictures, recordings, computerized cameras, PCs, and programming, which made the creation of child pornography moderately modest and furthermore, the advancement of the Internet innovation, which has expanded simplicity of creation and circulation of this material to astonishing statures.

Child pornography is progressively predominant in the present society and is currently one of the quickest developing internet exercises. In contrast to makers, holders of child pornography do not effectively take part in the physical and sexual maltreatment of kids. Notwithstanding, holders are watchers of this reported maltreatment and assault and can be, accordingly, likewise answerable for the interminable exploitation of guiltless youth. The far reaching accessibility of sex entertainment on the internet has worked up a sentimental frenzy shared by the legislature.

There have been numerous endeavors to constrain the accessibility of explicit substance on the internet by governments and law requirement bodies all around the globe. Sociological hypotheses of freak conduct have not been efficiently applied to the issue of who uses and who does not utilize digital erotic entertainment on the Internet.

Making laws is not equivalent to implementing them. While adequate enemy of child pornography laws exist in numerous countries, authorization is feeble. Besides, policing a worldwide activity like the internet includes policing residents from nations with generally contrasting local laws, societies and social mores. Despite the fact that these hindrances seem unfavorable, it must be recalled that child pornography is a substantive and convincing issue on worldwide, national, and nearby levels, and it is anything but an innocuous wrongdoing. Child pornography which involves child sexual abuse creates negative aspects in health of a child. Most of them face many health issues such as genital pain, genital bleedings etc. Most of the victims of child sexual abuse will have negative impact on their neurodevelopment and physical health and this also amount to long lasting depression [1].

A more noteworthy number of child molesters are presently utilizing PC innovation to arrange, keep up, and increment the size of their child pornographic assortments. Personally-manufactured unlawful pictures of children are particularly significant on the internet, and customarily molesters will exchange pictures of their sexual endeavors. At the point when these pictures arrive at the internet, they are lost and can keep on flowing always; in this way, the child is re-victimized as the pictures are seen over and over.

The quick development of the web and innovation has brought about the ascent and accessibility of child pornography in India. Considering these mechanical progressions and something else, the Indian Government has authorized different changes to fortify the legitimate structures.

India had hindered around 857 explicit sites in 2015 in view of the worries about child pornography. This specific choice was taken under the Information Technology Act and in consonance with Article 19(2) of the Constitution of India that permits the legislature to force limitations on the grounds of conventionality and profound quality. In any case, this total boycott was later lifted and just executed to those sites containing child pornography. As of late, again with the Department of Telecom has prohibited 827 destinations because of unlawful substance on sites. Current issues in the field of protection of children's are closely associated with rise in accessibility and usage of information and communication technologies (ICTs). The use of technology and the accessibility to internet services have increased worldwide. There must be an a urgent need to address the upcoming threat related to technology which impose new risks for the exploitation and abuse of children, as culprits misuse modern communication technologies to facilitate child sexual abuse. Therefore, states have to develop tailor-made instruments to tackle the specific dangers related to the use of ICTs by children [2].

No nation is invulnerable from child pornography, and it will require a deliberate exertion from governments, law implementation, and common society to guarantee that the world's children are ensured. India has attempted that approach and criminalization of a total prohibition on sex entertainment embrace.

2. Definition of child pornography

The legitimate age at which an individual can agree to sexual action fluctuates from nation to nation, a moving snag to the reliable and orchestrated assurance of children from sexual misuse on the universal level. While an individual younger than 18 might have the option to openly agree to sexual relations, such an individual is not lawfully ready to agree to any type of sexual abuse, including child pornography.

Besides, in conditions that require double guiltiness – when a wrongdoing carried out abroad should likewise be a wrongdoing in a wrongdoer's nation of origin for the guilty party to be arraigned in his/her nation of origin – concurrence on a typical age for what is a child is pivotal. Any inconsistency could forestall a child sex guilty party from being arraigned.

Hence, children, for reasons for child pornography enactment, ought to be characterized as anybody younger than 18 years.

Child pornography is distributing and transmitting vulgar material of children in electronic structure. As of late child pornography has expanded because of the simple access of the web, and effectively accessible recordings on the web.

Child pornography may incorporate genuine or recreated sex including minors, freak sexual acts, inhumanity, masturbation, sadomasochistic maltreatment, or the display of private parts in an explicitly exciting manner.

Child pornography is characterized by the Optional Protocol on the Sale of Children, Child prostitution and Child pornography as any portrayal of a child occupied with genuine or mimicked express sexual exercises or of the sexual pieces of a child for essentially sexual purposes.

Child pornography is the proof of the sexual maltreatment of a children and the creation of child pornography consistently surmises a wrongdoing submitted towards the child.

Child pornography typifies and debases youngsters.

Child pornography perhaps utilized by abusers as a way to control a child by guaranteeing that what is befalling the youngster in the image is something that numerous children participate in.

Child pornography can bring down the potential culprit's restraints and permits the wrongdoer to limit and misshape oppressive conduct. The culprit may utilize it as a support of his damaging conduct.

Offender's use of Child Pornography.

Guilty parties use child pornography for some reasons. Five of the most widely recognized include:

- Create a perpetual record for excitement and satisfaction.
- Lower child's restraints.
- Validate and affirm the child sex guilty party's conviction frameworks.
- Blackmail unfortunate casualties and other co-guilty parties.
- Sell for benefit or exchange.

3. Child pornography in India: general provisions

Sexual maltreatment among children in India has become wildly throughout the years, and an ongoing report by the Ministry of Women and Child Development expressing that over half of children have been manhandled comes as an eye-opener. Sexual maltreatment of children has not been another wonder, yet has won in the public eye for quite a while. Notwithstanding, the endeavors to check this marvel have been negligible, prompting an ascent in child sexual maltreatment.

Considerably after rehashed request by different partners to sanction another law to ensure children, such requests failed to be noticed. At last, the Government of India in the wake of setting up a draft Bill in the year 2006 passed the Protection of Children from Sexual Offenses Act, 2012. This exceptional enactment guarantees the assurance of children from sexual offenses lastly takes into consideration stricter discipline for such pedophiles.

There are different laws in India to secure and advance the offspring of the nation. In the Constitution itself, Article 21 accommodates the privilege to life and freedom, Article 24 does not permit children beneath 14 years to work in a mine, plant or take part in dangerous business. Article 39(f) makes it required for the State to coordinate its approach towards making sure about the wellbeing and quality of children and to give those openings and offices to grow steadily and Article 45 gives that the State will attempt to give youth care and training to children beneath the age of 6 years [3]. There likewise exist uncommon laws for violations against children, for example, The Immoral Traffic (Prevention) Act, 1986, The Child Marriage Restraint Act, The Child Labor (Prohibition and Regulation) Act, 1986 and The Juvenile Justice (Care and Protection of Children) Act, 2000.

The Penal Code, 1860 and The Criminal Procedure Code, 1973 oversees the substantive and procedural pieces of criminal offenses, including those which apply to children. Since no uncommon arrangements are overseeing the maltreatment of children, similar laws apply to the grown-ups and offspring of the nation. The laws overseeing sexual offenses incorporate Sections 375 (Rape), 377 (unnatural offenses) and 354 (shocking the humility of ladies) under the Penal Code. There are additionally offenses against minor young ladies for example Section 372 (selling of young ladies for prostitution) and Section 373 (purchasing of young ladies for prostitution) [4]. In any case, these laws are not exhaustive or satisfactory to deal with such grave offenses on such delicate matured children. These arrangements are additionally one-sided towards ladies and are insufficient themselves either substantively or procedurally to meet the exceptional needs of sexual maltreatment among children.

Disregarding such broad laws, the State of Goa passed the Goa Children's Act, 2003 to ensure, advance and safeguard the interests of children in Goa and to make a general public that is pleased to be child inviting. The demonstration isolates the offenses into grave rape which covers various kinds of intercourse-vaginal, oral, butt-centric, utilization of articles, constraining minors to have intercourse with one another, purposely making injury the sexual organs and making children present for explicit photographs or movies; rape which covers sexual contacting with the utilization of anyone part or item, voyeurism, exhibitionism, indicating obscene pictures of movies to minors, making children watch others occupied with sexual action, giving of dangers to explicitly mishandling a minor, loudly manhandling a minor utilizing revolting and profane language; and interbreeding which is the commission of a sexual offense by a grown-up or a child who is a relative through ties of appropriation. In this manner, this was the main enactment of India constrained to Goa, where there were exceptional laws to shield children from sexual maltreatment.

The absence of satisfactory laws was likewise referenced in different cases under the watchful eye of the Supreme Court of India. In India the applicants needed the intra-State dealing of small children, their subjugation and coercive repressions, customary inappropriate behavior and maltreatment to be made cognizable under the Indian Penal Code. The Supreme Court of India likewise made a referral to the Law Commission of India on issues of child sexual maltreatment.

The Law Commission expressed that the instances of penile entrance were secured under Section 375, the unnatural offenses, for example, fleshly intercourse against the request for nature with any man, lady or creature were taken consideration by Section 377 and the infiltration of finger or lifeless thing into the vagina or rear-end against the desire of a lady or female child would be secured by Section 354.

The avocation given by the Law Commission was that the gravity of these different offenses were extraordinary and in this way, the offenses referenced under Sections 354 and 377 ought not be brought under the proviso of assault or be given such unforgiving discipline and consequently there was no compelling reason to carry any new law into the image. In any case, one feels that all the previously mentioned offenses are egregious and there ought to be stricter discipline forced on such guilty parties.

Without stricter rules for unfortunate casualty security, the Supreme Court itself detailed different rules for the assault injured individual. The court expressed that because of the actuation of extraordinary dread or because of the stunned State of the person in question; the injured individual will most likely be unable to give full subtleties of the episode, which may prompt an unnatural birth cycle of equity. The inquiries in this way presented to the unfortunate casualty in court may prompt shame of the person in question, because of which an injured individual may not be agreeable, and subsequently, the Court asked the Presiding Officer instead of the restricting direction to offer the pertinent conversation starters to the person in question.

The Court additionally requested that the exploited people be permitted breaks and adequate time to respond to the inquiries. The Court additionally proposed holding such preliminaries in the camera, to make the unfortunate casualty progressively agreeable, and to guarantee that the injured individual can respond to the inquiries effortlessly, thus that the injured individual is not reluctant and is coming clean.

Another token of our lacking laws is the Report of the National Crime Records Bureau concerning child sexual maltreatment. The records show that a sum of 5484 kid assault cases were accounted for during the year 2010, an expansion from 5368 in the year 2009, 679 instances of procuration of minor young ladies were accounted for in 2000 against 237 out of 2009. Seventy-eight instances of purchasing young ladies and 130 instances of selling of young ladies for prostitution were accounted for in the year 2010 against 32 and 57 of every 2009.

The investigation of Child Abuse by the Government of India in the year 2007 gave some stunning disclosures. It was discovered that 53.22% of youngsters had confronted at least one types of sexual maltreatment and half of such maltreatment were from people known to the child or were people in a place of trust and obligation.

In the light of the grave circumstance confronting children in India, today, the Protection of Children from Sexual Offenses Bill was made in the year 2006 and was at long last passed by the Indian Parliament in 2012.

4. Protection of children from sexual offenses act, 2012

The Protection of Children from Sexual Offenses Act, 2012 was ordered with the article to shield the children from offenses of rape, inappropriate behavior, sex entertainment and to accommodate the foundation of Special Courts for the preliminary of such offenses and matters associated therewith or accidental thereto. The Act gets its capacity from Article 15(3) of the Constitution of India, which enables the State to make uncommon arrangements concerning children. Article 39(f) of the Constitution of India accommodates the State to guide its approach to make sure about the children with the goal that they are not manhandled and their adolescence and youth are ensured against abuse. The State additionally means to satisfy its acknowledgment of the Convention on the Rights of Child, which was consented by India on 11-12-1992.

The Convention basically features the measures that should be embraced by the State to forestall:

1. Affectation or compulsion of a kid to take part in any unlawful sexual action,
2. The exploitative utilization of youngsters in prostitution or other unlawful sexual practices,
3. The exploitative utilization of youngsters in explicit exhibitions and materials.

Simultaneously, the Act expects to guarantee the best possible improvement of the children and means to secure their protection and privacy through the legal procedure and to guarantee the physical, passionate, scholarly and social advancement of the child.

The using of child for pornographic purposes is an offense.

The term here implies that the use of the child in any such type of media including system or notice by TV slots, web or some other electronic structure or the printed structure which could possibly be for individual use or appropriation might be an offense in the event that it is utilized for sexual satisfaction.

This incorporates the portrayal of sexual organs, use of a child in genuine or reproduced sexual acts or foul or profane portrayal of a child. The Act orders a discipline of a limit of 5 years and in the second conviction this may stretch out as long as 7 years with a fine. In the event that the individual additionally participates in such a demonstration, that establishes the previously mentioned sexual acts/attacks, he would be subject forever detainment.

Additionally, an individual who stores obscene material for business purposes in any structure including a child will be rebuffed with detainment reaching out as long as 3 years or fine or both. The enactment not just rebuffs the wrongdoer who submits such acts yet.

Additionally people who abet or endeavor to submit such a demonstration. An individual who abets the commission of the offense by actuating, planning, deliberately supporting by any demonstration or oversight would be at risk for the offense and would be culpable for as long as 1 year or with fine or both.

The Act is by all accounts an extensive bit of enactment. The Act begins from characterizing the different offenses to rebuffing people abetting such an offense. The feature of the Act is the methodology and the shields intended to ensure and cause the children to feel safe so that there is negligible long haul sway because of the terrible wrongdoing.

5. International laws

5.1 International aspects related to child pornography

Child pornography is a multi-jurisdictional issue to which a worldwide methodology must be applied. Effectively fighting child pornography and child misuse on a worldwide scale requires uniform enactment; laws that fluctuate from nation to nation serve to debilitate the position against child sexual abuse and permit child predators to a mass endeavors in nations where they realize they are best ready to abuse children. A comprehensive and uniform methodology is the best methods for battling the sexual abuse of children since it takes into consideration consistency in criminalization and discipline, it raises open consciousness of the issue, it expands administrations accessible to help unfortunate casualties, and it improves in general law requirement endeavors at the national and worldwide levels. Agreeing to global legitimate models is an underlying advance in tending to child pornography, to be trailed by national executing enactment and the production of a national administrative plan to battle child sex entertainment.

Under International law the first convention which was implemented on children's right was (CRC) convention of rights of children. This convention lays the guidelines on right enjoyed by the children's and also give accurate meaning to the term child. It is comprehensive convention which deals with every aspect of a child right. Article 2 of the convention make it clear that there must not be any discrimination made in respect to race, language, sex, religion or any other status when you are empowering the child with their rights.

Article 19(1) states that every child need to be protected from different type of physical and mental violence specifically sexual maltreatment, sexual abuse and sexual exploitation. This amounts to be most essential article to put end to sexual exploitation of children's. Article 23 of this convention is concerned with right of children with disabilities. Other than convention on rights of children we have governed with one declaration which also deals with protection of children's that is universal Declaration of human rights under article 25(2) of the declaration it is been observed that every children born out or without wedlock needs to be equally protected. Likewise in International covenant on civil and political right we have Article 24(1) which states that every child will be protected without any discrimination in respected of race, sex, color accordingly as it is required y his status as a minor on the part of his family and society [5].

6. Comparative analysis of legislations of the United Kingdom, India and South Africa

The United Kingdom passed its enactment for children in the year 2003, comparably, the revised enactment of South Africa was spent in the year 2007 lastly, and the Indian enactment was spent in the year 2012. Since the Indian enactment was framed in the wake of investigating the United Kingdom and South African enactment, a look, and examination with the parent institution would assist us with dissecting the deviation and contrasts between these Acts.

In the United Kingdom, (UK) this enactment is called as Sexual Offenses Act, 2003, in India, The Protection of Children from Sexual Offenses Act, 2012 and in South Africa Criminal Law (Sexual Offenses and Related Matters) Amendment Act, 2007.

The Act to ensure children in the UK was enacted with the article to forestall and shield the children from hurt from sexual acts. The South African Act accentuates the need to address the helplessness of children and furthermore features the social marvel of children misuse which looks to make the general public broken though the Indian Act was pressing enactment authorized to satisfy the need of great importance to check and forestall an expansion in the quantity of child sex misuse cases.

In the UK, an offense on a child beneath 13 is severally rebuffed and the other class for children is 16 years for example for genuine sexual offenses. In South Africa, children are characterized as those being under 18, be that as it may, another classification has likewise been made wherein, children somewhere in the range of 12–16 years, on the off chance that they enjoy sexual exercises with one another, both might be arraigned with the consent of the important position. In India, the separation is concerning the time of assent. In specific cases, wherein the child is somewhere in the range of 16–18 years old, the Court would try to discover whether there was assent between the child and the grown-up or not.

The UK Act additionally remembers attack for a child younger than 13 by infiltration, without the utilization of the penis, yet with the utilization of any piece of the body or any such item. Both these offenses would prompt life detainment. In the UK, the age-furthest reaches of 13 are of criticalness, as any such offense on a child beneath 13 is met with graver discipline. Regardless of whether an individual impels a child for example beneath 16 to participate in sexual movement prompting infiltration, the individual would be subject to a most extreme detainment of 14 years.

So also, any such infiltration of a child underneath 16, of the butt, vagina or mouth with the penis or any piece of the body or whatever else, would be at risk to a term not surpassing 14 years. The prompting or making a child participate in such movement is additionally.

An offense, henceforth, it is abundantly certain that in the United Kingdom, as far as possible to be considered as a child is 16 years. There is no understanding of assent, in any case, it would be seen whether the individual realized that the child was underneath 16 years old.

In India, be that as it may, rather than assault the terms utilized are penetrative rape and bothered penetrative rape. These demonstrations incorporate the infiltration by a penis, anyone part or article, which might be finished by the individual, or the child on the individual. The age of the children ought to be underneath 16, and the Court would see whether assent was given if the child is somewhere in the range of 16–18 years old. The discipline for this offense is at least 7 years, which may stretch out to life detainment. On account of people in power, the base discipline is 10 years and the most extreme being life detainment. Consequently, the age of the child is viewed as beneath 18 years, and for an assent somewhere in the range of 16–18 years.

In the South African Act, the term Sexual Activity incorporates sexual infiltration and it goes under the head sexual abuse of child. There are different sorts of offenses referenced right now. These arrangements rebuff an individual who takes part in the administrations of a child, with or without his assent, or when the administration is offered to a third individual or an individual who permits the commission of such an offense, or who gets a compensation for the sexual demonstration with the child is rebuffed under the Act. The arrangement likewise rebuffs an individual, who makes travel game plans for or for the benefit of the third individual to encourage the commission of a sexual demonstration. Henceforth, in the South African Act, assent is unimportant, and children are the individuals who are underneath the age of 18. In any case, for assent, as far as possible is between 16 and 18 years.

In the United Kingdom, taking part in sexual action within the sight of a child is an offense with a discipline of 10 years, while making a child watch a sexual demonstration is likewise an offense. Notwithstanding the abovementioned, inducing or making a child be associated with sex entertainment is likewise an offense. In India, be that as it may, demonstrating the child any article in any structure for explicit designs is named as Sexual Harassment. The utilizing of children in any type of media, for sexual delight, which may incorporate portrayal of sexual organs, connecting with the kid in genuine or reproduced acts or the profane or indecent portrayal of child is an offense. In South Africa, the presentation or show of child pornography or sex entertainment or convincing or making the kids observer sexual offenses, sexual goes about just as self-masturbation is an offense. The utilization of children for or to profit by child pornography is an offense. Hence, in each of the three nations, the utilization of a child to take part in sexual action or to make him observe any sexual demonstration is an offense with discipline.

Regardless of whether an individual supplies, uncovered or shows an article to be utilized for a sexual demonstration, child pornography, distribution or film would be at risk. The arrangement additionally incorporates any course of action that might be done in any piece of the world, or when an individual welcomes, convinces, lures, prompts or forces a child to travel abroad, or cause for a gathering to be held, for the commission of the sexual.

Demonstration would be subject for the offense of sexual preparing. To put it plainly, it implies the abetment of sexual maltreatment to children. In the UK, abetment has been utilized in an alternate arrangement, wherein, when an individual orchestrates or encourages the commission of a child sex offense, he would be at risk. In India, the term utilized is abetment, which includes the affectation, connecting with at least one people or purposefully helping an individual to submit an offense.

One of the significant segments, which exist planning to shield the children from maltreatment from individuals in a place of trust, is available in both the Indian just as the UK Act. A demonstration of sexual nature or even a demonstration to cause, to actuate a child to take part in sexual movement, to make a child watch a sexual demonstration or to do any sexual demonstration within the sight of a child is an offense.

The individuals are said to be in a place of trust when the children realizes that such people are people in the situation of trust. The places of trust are characterized as the position where the individual cares for the child kept in a foundation, in a medical clinic, autonomous center, a consideration home, regardless of whether private or not, a network home. The situation of trust additionally incorporates people getting or not accepting training, where the child is getting instruction, an individual who prompts, minds or oversees the children. In India, there are terms, for example, exasperated penetrative rape or penetrative rape, wherein the individual in places of trust, for example, cop, individual from military or security powers, a community worker, an individual who deals with an emergency clinic or instructive organization, when submits a sexual demonstration, he would be at risk for discipline. There are no such arrangements concerning the situation of maltreatment in South Africa.

In the present world, where there has been an expansion in the abuse of positions by specialists, it is astonishing to see that the South African Act does not take a harsh remain on this issue. It gets one of the significant arrangements, since it is additionally observed that such people may utilize their impact to conceal such cases.

The UK and the Indian Act both rebuff a relative who submits a sexual offense on a child. A relative is the individual who might be a parent, grandparent, sibling, sister, stepbrother, stepsister, auntie or uncle, temporary parent, step-parent, cousins, step-sibling or sister, who lives in a similar family unit and is routinely associated with thinking about, preparing, managing or being the sole in-control. In India, a relative is characterized a relative of the youngster through blood or reception or marriage or guardianship or in child care, or having a residential relationship with a parent of the child, or who are living in the equivalent or imparted the family unit to the child.

In conclusion, the investigation of the three establishments would show that all the three enactments have been detailed by the cultural needs of the nation. The UK Act is by all accounts the most exhaustive as it traces everything and has a different arrangement for a wide range of acts and offenses. The discipline for the offenses is likewise stricter than the other two nations. Then again, the time of child is taken to be underneath 16 years. This again changes from one nation to the next.

The South African Act underscores more on the activities of the third people or acts which expect to encourage the commission of the offense. There is a nonattendance of any arrangement concerning an individual in the situation of obligation or authority. This ought to have been incorporated since, in a nation like South Africa, there would be more maltreatment of intensity. The age of the child is underneath 18 years, yet the period of assent is somewhere in the range of 12–16 years.

In India, the Act aims to join the over two enactments to the degree conceivable, including the progressions that may be required according to the necessities of the general public. A child is characterized as being underneath 18 years, yet the time of assent is between 16 and 18 years. There are no arrangements concerning preparing or making a trip to submit a sexual demonstration.

7. Conclusion

The legitimate and procedural boundaries to ensuring the interests of children on the Internet are vexing. Definitional challenges, just as various social and social mores, make troubles concerning contriving a compelling global structure for ensuring child on the web. The issues are additionally exacerbated by various methodologies that have been received concerning issues including the effort of criminal ward over exercises led by means of the vehicle of the internet, removal and the acquiring of proof. The absence of a steady and amicable structure on security, content guideline, and pornography additionally goes about as significant obstructions to affecting a serviceable worldwide technique to ensure the interests of children on the internet. Be that as it may, as the conversation likewise looks to the show, the challenges are not impossible.

There must be an assurance with respect to all nations to secure children that a successful legitimate system would then be able to be conceived. With a level of legal resourcefulness in receiving a wide perusing of existing offense-making arrangements in existing criminal rules, auspicious administrative intercession to fill in the escape clauses and a reasonable level of purposeful worldwide co-activity in the field much should be possible in the continuous the fight to secure children. It is maybe able to end the conversation by repeating a statement that gets straight to the point and helps put the issues in context that a child's life is unmistakably increasingly significant then those sorts of moderately minor worries about common freedoms and capture. Those are significant inquiries however set them against a child's life, a kid's mental prosperity, and in all honesty, there is just a single conceivable answer.

Over the years, different research in regards to the status of child pornography enactment around the globe has shown that gradual advancement is being made. Different global legitimate instruments are set up, which has helped bring issues to light and connect new criticalness to this reason. It stays clear, in any case, that more nations need to make a move now on the off chance that we are to make sure about a more secure future for the world's children. While fighting child pornography at home and abroad is an overwhelming assignment, harmonization of laws is fundamental to viably address this developing, worldwide marvel. Building up an exact meaning of the term "indecency" is troublesome. What might be considered as vulgar in one nation may not be considered as disgusting in another. It chiefly relies upon the good and moral estimations of the individuals who have a place with a particular nation. In any case, the nonexclusive meaning of vulgarity alludes to a demonstration or discourse or thing that is probably going to degenerate the profound quality of the overall population as a result of its foulness or lasciviousness in substance or structure.

The show of something hostile to humility or goodness or articulation of unchaste or lecherous thoughts or being disgusting or indecent is viewed as vulgar, in many nations. As I would see it, to control child pornography, we ought to totally boycott pornography destinations. This stringent activity can take care of the issue to a bigger degree. This would give us an opportunity to think and plan some better approaches to annihilate child pornography around the world. Delineation of minors, both genuine and virtual, just as grown-ups giving off an impression of being minors, in electronic child pornography, ought to be forestalled by Indian law. Stringent estimates must be taken to battle such egregious maltreatment.

Author details

Deepa Salian* and Sofia Khatun
Christ Deemed to be University, Bengaluru, India

*Address all correspondence to: deepa81salian@yahoo.com

References

[1] Pg-5 Psychological Society [Internet]. 2013. Available from: https://www.google.com/search=www.psychology.org.an%2fabout-us%2fpaper+discussion+paper+review%2fchild+sexual+abuse-review+paper.&aqs=chrome..69i57j69i58.800j0j7&sourceid=chrome&ie=utf-8 [Accessed: 23 April 2020]

[2] Pg-3 Child Online Sexual Abuse [Internet]. 2017 Available from: WWW.UNICEF.ORG/NAMBIA/NA.COP-LEGAL-ANALYSIS PDF [Accessed: 23 April 2020]

[3] Pal R, Pal S. M.P. Jain Indian Constitutional Law. 6th ed. Nagpur: Lexisnexis Butterworth Wadhwa; 2010. pp. 1503-1504. ISBN: 978-81-8038-621-3

[4] Thomas KT, Rashid MA. The Indian Penal Code. 34th ed. Haryana: LexisNexis; 2014. pp. 857, 861, 928. ISBN: 978-93-5143-099-5

[5] Pg-1 Right of child [Internet]. 2003. Available from: un.org/esa/socdev/enable/comp501.htm [Accessed: 01 May 2020]

Section 3

Digital Evidence

Chapter 5

Novel Methods for Forensic Multimedia Data Analysis: Part I

Petra Perner

Abstract

The increased usage of digital media in daily life has resulted in the demand for novel multimedia data analysis techniques that can help to use these data for forensic purposes. Processing of such data for police investigation and as evidence in a court of law, such that data interpretation is reliable, trustworthy, and efficient in terms of human time and other resources required, will help greatly to speed up investigation and make investigation more effective. If such data are to be used as evidence in a court of law, techniques that can confirm origin and integrity are necessary. In this chapter, we are proposing a new concept for new multimedia processing techniques for varied multimedia sources. We describe the background and motivation for our work. The overall system architecture is explained. We present the data to be used. After a review of the state of the art of related work of the multimedia data we consider in this work, we describe the method and techniques we are developing that go beyond the state of the art. The work will be continued in a Chapter Part II of this topic.

Keywords: multimedia forensic data analysis, standardization of forensic data analysis, video and image enhancement, video analysis, image analysis, speech analysis, case-based reasoning, multimedia feature extraction, handwriting, Twitter data analysis, novelty detection, legal aspects

1. Introduction

The objective of this work is to provide novel methods and techniques for the analysis of forensic multimedia data. These methods and techniques should form a novel toolkit for automatic forensic multimedia data. The data modalities the proposed work is considering are images and videos, text, handwriting, speech and audio signals, social media data, log data, and genetic data. The integration of methods for all these different data modalities in one tool kit should allow the cross-analysis of these data and the detection of events by interlinking between these data. The proposed methods will face on standard forensic tasks, for example, identification of events, persons, or groups and device recognition. Together with the end users and the police forces, new standard tasks will be worked out during the project and will give a new input to the standardization aspect of forensic data analysis.

The proposed novel methods and techniques will consider all aspects of multimedia data analysis such as device identification and trustworthiness of the data, signal enhancement, preprocessing, feature extraction, signal and data analysis, and interpretation.

Techniques for detecting artifacts in images and videos are of paramount importance. To trust the information extracted from images and videos, it is necessary to make sure that the image and video have been recorded by a camera, and that no artifact has been added. The detection of artifacts is a key element to use an image or a video in court. Thus, it should be clearly assessed the integrity of images and videos used as a proof of evidence.

In most image applications, the acquired images represent a degraded version of the original scene. Degradation in such images may appear in different forms. These types of degradations must be removed before the images are used for classification or decision making.

Novelty detection for the identification of novel situation and tasks will be another task that will be important in forensic applications, where the victims or events are very flexible. It will allow to identify new tasks, and by doing so, it will be an automatic method to improve standardization of the analysis of forensic data.

We will also develop learning methods to include new data into the existing cases and summarization of new and old cases into more general cases applicable to a wider range of tasks for further law purposes. For that, novel case-based reasoning methods will be developed that can keep the cases based on their multimedia features and specific event features in a case base, so that they can be easily retrieved and applied for new situations. The case-based reasoning system will consist of novel probabilistic and similarity-based methods. It will provide a wide range of novel similarity-based reasoning methods for the different feature types for identification and similarity determination. A special taxonomy for similarity determination and measures will be worked out and implemented in the CBR system. It will provide explanation capabilities for similarity and as those it will help a forensic data analyst to identify the right reasoning method for his particular problem. This aspect goes along with the training and education aspect for forensic data analysis. Part of this will be self-contained in the chosen methods and realized by the system.

In Section 2, the background and the motivation of our work will be described. Taking into account the special needs for multimedia forensic analysis, identification, and recognition system, we develop a novel architecture based on case-based reasoning. The data used are described in Section 3. Related work and the progress we want to make with our work are described in Section 4. This work does not only take into account to develop novel methods and techniques for multimedia content processing and reasoning, but we are also taking into account the legal aspect that is going along with processing sensible data. Finally, we given conclusions in Section 5. This chapter is continued in the Chapter Part II of Novel Methods for Forensic Multimedia Data Analysis.

2. Background, motivation, and overall system architecture

The analysis of multimedia data has to consider different aspects of the modalities of the data. We want to deal with images and videos, text, handwriting, speech and audio signals, social media data, log data, and genetic data. The idea is to come up with an automatic system that should cover all aspects of data analysis for the different modalities from the signal enhancement, preprocessing, feature extraction to the analysis, and interpretation. This includes image enhancement in order to eliminate the degradation in an image that might appear because of a known or an unknown blurring function, which leads to the consideration of deconvolution and blind deconvolution problems or because of very low resolution devices, which lead to the combination of several low resolution images to obtain a high resolution one, the so called, super-resolution problem or to the utilization of highly compressed images, which suffer from compression artifacts.

Techniques for detecting artifacts in images and videos will be developed to trust the information extracted from images and videos. They should allow to make sure that the image and video have been recorded by a camera, and that no artifact has been added.

Feature extraction will be the selection of a set of sufficiently low- and high-level features in order to complement the existing standards for image, video, and audio data, with the aim at enabling novel and robust classification and recognition methods. They should allow modeling the standard tasks for forensic data analysis known so far but should be flexible enough to cover the needs of newly arising task.

Twitter was actively used by rivaling gang members to plan their assaults. Twitter data are hard to analyze because the text fragments are very short, multiple persons can be involved in a conversation about various topics, and the data are rapidly changing. Novel methods are necessary, which can be used to monitor in real-time Twitter and identify potential threats including individuals and communities of users who are planning illegal activities.

Furthermore, we plan to build a dynamic model on Twitter text to forecast the upcoming significant events and emotions of the crowd associated with these events. While there can be many events with strong presence in Social Media, some of them would have stronger negative emotions associated with them. These events are candidates that may have criminal nature or significant social consequences.

The huge amount of CCTV systems has increased the importance of video and image evidence in forensic labs. An automatic system should be able to select heads, vehicles, license plates, guns, dresses, and all other objects that can link a person to the event.

An important main focus of police work is the identification of people for which a decision of the public prosecutor's office or a judge to the observation or an arrest warrant was issued. Within the scope of this arrangement, the use of video supervised places and facilities, or at before not known places, the application of mobile video technology should occur for this purpose. The aim is to develop methods and procedures for an automatic system for identification of one or several target people in mobile video recordings based on passport photos or other available pictures.

A significant portion of data collected by Law Enforcement Agencies consists of speech and audio files. They form an important part of legal cases. Speech recognition systems (such as dictation systems) are now available in many languages. However, continuous spontaneous speech recognition is still an unsolved problem. Novel methods for the recognition of continuous spontaneous speech and other audio signals are necessary.

While the commercially available optical character recognition systems are very successful for printed documents, recognition of words in unconstrained settings or "in the wild" still is an open problem, and recognition of handwritten text continues to be a challenge. We propose to develop novel Handwriting Recognition Methods for unconstrained settings.

Novel Case-Based Reasoning (CBR) methods will be developed for the recognition, interpretation, and identification task. Case-based reasoning explicitly uses past cases from the domain expert's successful or failing experiences. CBR is very useful in applications, where generalized knowledge is lacking. Therefore, case-based reasoning can be seen as a method for problem solving as well as a

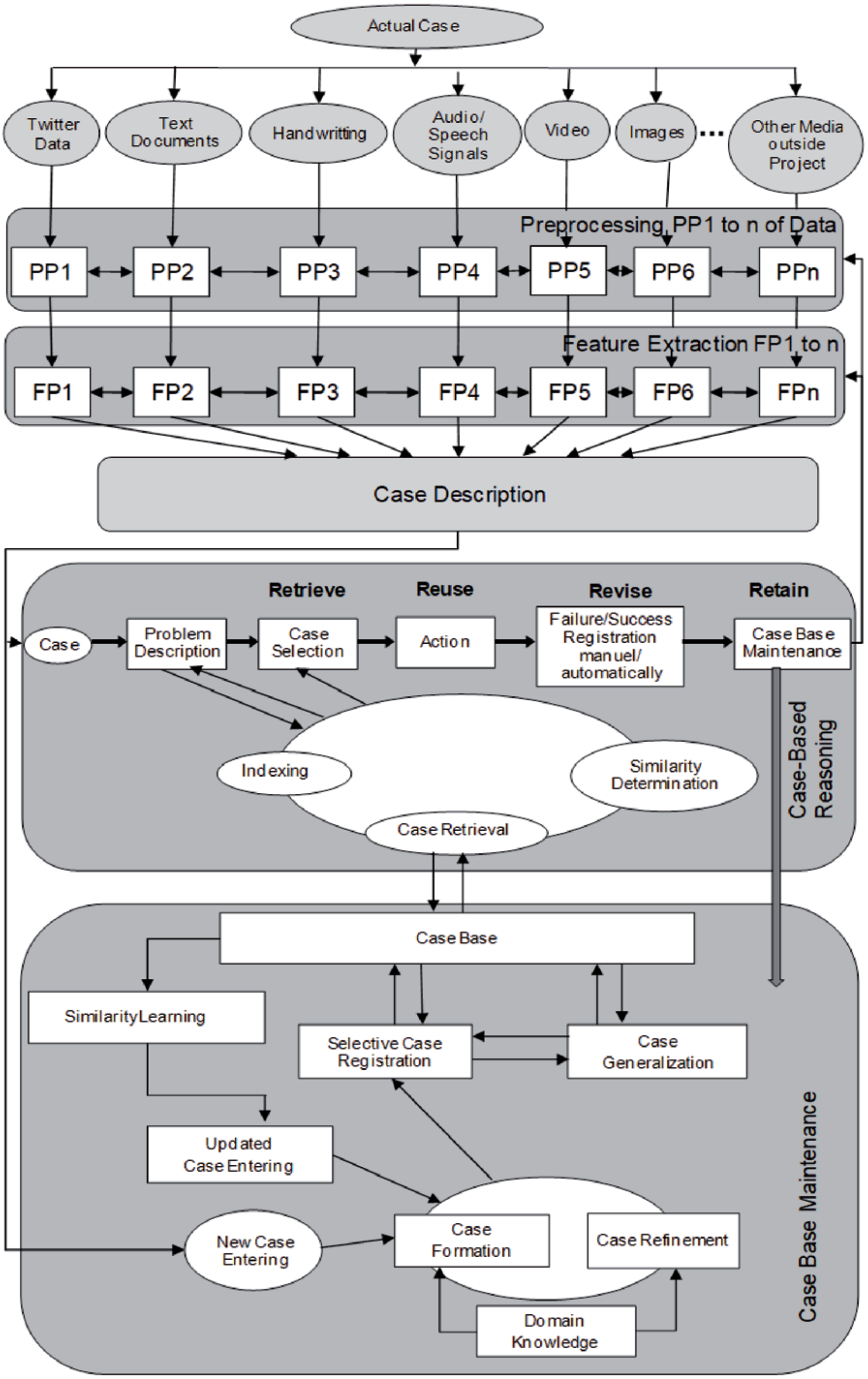


Figure 1.
System overview.

method to capture new experiences and make them immediately available for problem solving. It can be seen as a learning and knowledge discovery approach since it can capture from new experiences some general knowledge such as case classes, prototypes, and some higher-level concept. All these points make a CBR system very useful for the analyses of forensic data. The method is able to capture new cases and store new and old cases in a summarized way, so that they can be easily retrieved or used for reasoning. The reasoning methods are based on similarity that makes it very useful to detect and identify similar and identical cases without having generalized knowledge. Different similarity measures have to be developed that can deal with the different modalities of data and their case representation. A taxonomy of similarity will be developed that explains the relation, usefulness, and application of the different similarity measures to the data that will help a forensic data analyst to efficiently apply these reasoning methods to his problem.

All the above-mentioned facts result in the overall system architecture given in **Figure 1**. The architecture consists of the three main processing units: media preprocessing, feature extraction, and decision unit based on case-based reasoning. The input is the different media data. The architecture is open, so that new input media data can be considered when the necessary processing modules are available. The outcome of the preprocessing and the feature extraction unit is a description of the different media data by sufficiently low- and high-level features that will be combined to the case representation. The reasoning will be done by the case-based reasoning unit based on formerly calculated case representation. The reasoning will be the identification and recognition of the objects or scenario's as well as the detection of novel events. The CBR unit will be criticized based on the result of the action, and the decision of the CBR unit has been proposed. Depending on that outcome, case-based maintenance will be done. New case will be stored in the case base, the similarity measure will be updated or changed, or case generalization will be done.

Besides the development of novel processing and reasoning methods, it is necessary to develop a legal framework regulating the process of gathering, processing, analyzing, and integrating multimedia data.

3. Data used

Different types of security-related data will be used for the work provided by the end users:

- Passive millimeter-wave (PMMW) images and video are used for security screening as many materials, including clothing, are transparent to millimeter-waves. The imagers that use this technology, such as those developed by ALFA [1], are therefore installed at security checkpoints to screen people for hidden weapons (including powders, liquids, and gels) and contraband. They are characterized by a low resolution compared to visible images, due to the wavelength used. ALFA's current software automatically detects objects within the spatial and thermal resolution of the system and draws a red box around them. Some examples of this image type are given in **Figures 2–4**. These are then represented at the approximate locations on a generic silhouette to preserve the subject's privacy. However, object classification to automatically distinguish between a threat and a nonthreat object is not currently performed. A new system will be developed to make a classification based on the shape and size of the objects detected in the raw millimeter-wave image. This would reduce the number of false alarms.

- Anonymous Data from Text will be collected. These data are freely available on the Web. We propose to perform initial experiments on anonymized data to validate the feasibility of our approach. After authorization of the responsible superiors of the cybercrime unit is obtained, we will use the developed system for real-life investigations.

- A Telekom company will prepare a speech database obtained under various conditions and under various speech coders and encoders to test the new algorithms.

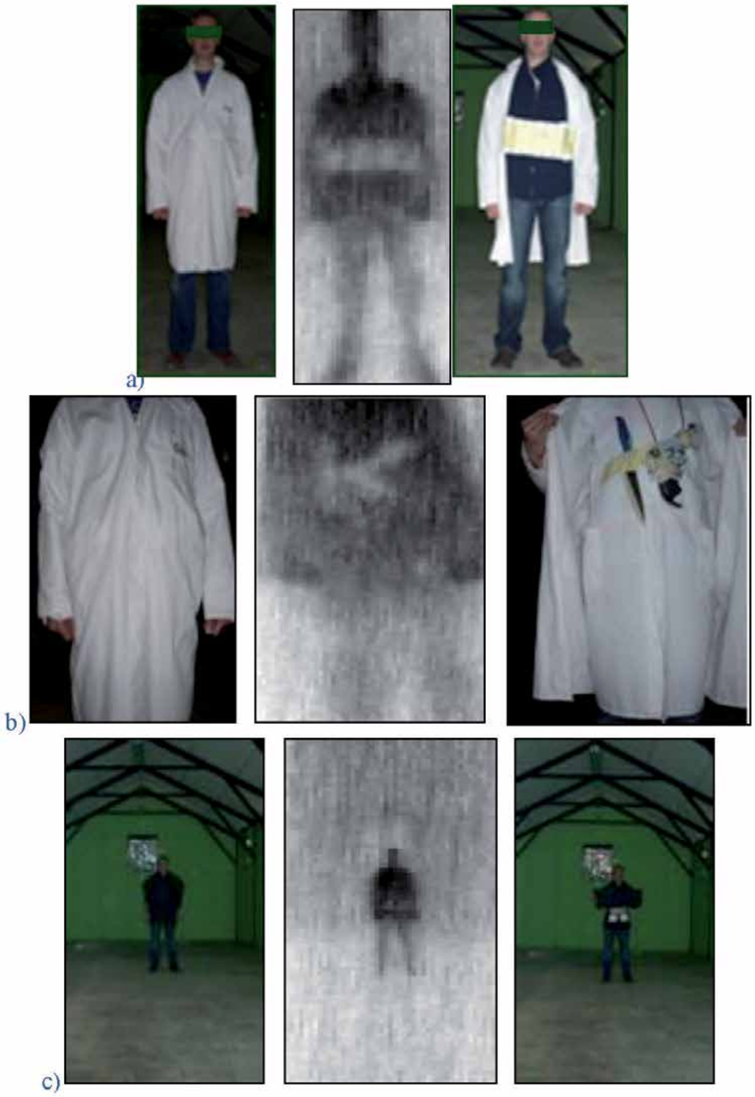


Figure 2.
(a) Left: clothed subject; center: raw millimeter-wave image of subject; right: subject showing hidden suicide bomber belt; (b) left: clothed subject; center: raw millimeter-wave image of subject; right: subject showing hidden gun and knife; (c) left: clothed subject at 10 m; center: millimeter-wave image of subject at 10 m; right: subject showing two hidden bags of powder explosives. Subject with gel pack hidden between the legs and automatic millimeter-wave detection marked + raw millimeter-wave image of subject; right: subject with gel pack hidden under the arm and automatic millimeter-wave detection marked + raw millimeter-wave image of subject.

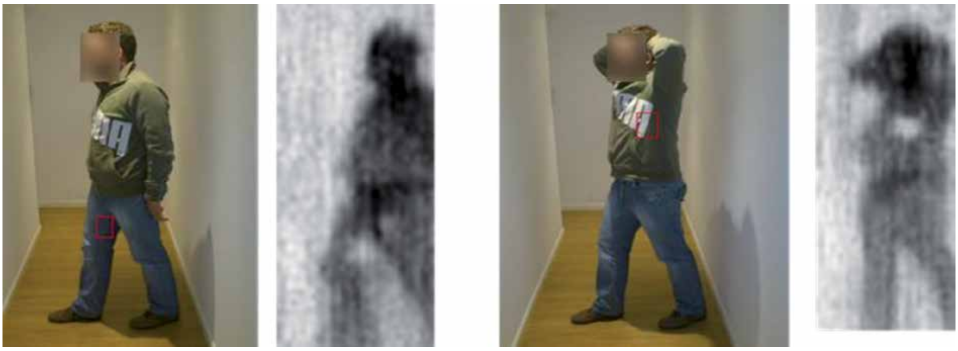


Figure 3.
Automatic object and potential threat detection (ATD) on processed millimeter-wave image on the left and privacy protection output to operator on the right.

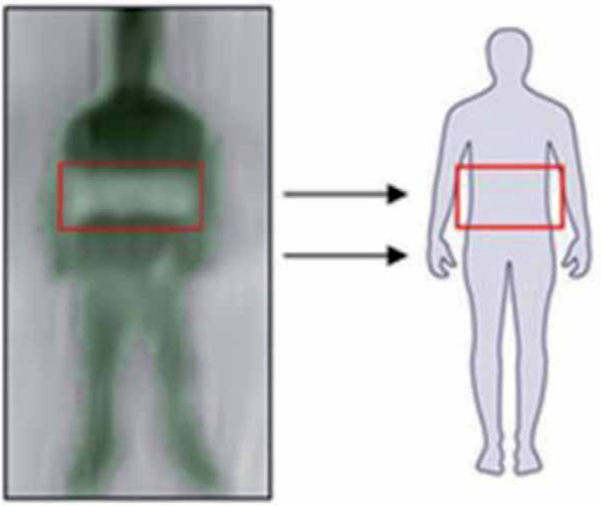


Figure 4.
Person with hidden object around the hip.

- Video and Image databases with case scenarios will be provided by police forces.

- Handwriting documents will be collected through the involvement of graduate and undergraduate students. We also plan to use the following benchmark data set: IAM Database for Off-line Cursive Handwritten Text http://www.iam.unibe.ch/~zimmerma/iamdb/iamdb.html. The database contains the forms of unconstrained western handwritten text. It includes 27,000 isolated words (400 pages).

4. Related work and progress

4.1 Video and image enhancement, filtering, and assessment

4.1.1 State of the art

In most image applications, the acquired images represent a degraded version of the original scene. These applications include astronomical imaging [2] (e.g., using ground-based imaging systems or extraterrestrial observations of the earth and the planets), commercial photography [3, 4], surveillance and forensics [5, 6], medical imaging [7] (e.g., X-rays, digital angiograms, autoradiographs, MRI, and SPECT), and security tasks where commercial photography and other image modalities like Synthetic Aperture Radar (SAR) [8] and Passive Millimeter (PMMW) [9] are frequently used.

Degradations in such images may appear in different forms. They may be due to a known or an unknown blurring function that leads to the consideration of deconvolution [9–13] and blind deconvolution [3, 14] problems. They may also be due to the use of very low-resolution devices, which lead to the combination of several low-resolution images to obtain a high-resolution one, the so called, super-resolution problem [15, 16] or to the utilization of highly compressed images, which suffer from compression artifacts [17]. These types of degradations must be removed before the images or video sequences are used for classification or decision making. Interestingly, all the problems described above can be formulated within the Bayesian framework [18–20]. A fundamental principle of the Bayesian philosophy is to regard all parameters and unobservable variables as unknown stochastic quantities, assigning probability distributions based on subjective beliefs. Thus, the original image(s), the observation noise, and even the function(s) defining the acquisition process are all treated as samples of random fields, with corresponding prior probability density functions that model our knowledge about the imaging process and the nature of images.

4.1.2 Beyond the state of the art

Once the problem is modeled, inference is then needed. The recently developed variational Bayesian methods have attracted a lot of interest in Bayesian statistics, machine learning, and related areas [18–20]. A major disadvantage of traditional methods (such as expectation maximization (EM)) is that they generally require exact knowledge of the posterior distributions of the unknowns, or poor approximations of them are used. Variational Bayesian methods overcome this limitation by approximating the unknown posterior distributions with simpler, analytically tractable distributions, which allow for the computation of the needed expectations and therefore extend the applicability of Bayesian inference to a much wider range of modeling options: more complex priors (which are very much needed in applications involving images) modeling the unknowns can be utilized with ease, resulting in improved estimation accuracy.

Techniques for detecting artifacts in images and videos are of paramount importance. In order to trust the information extracted from images and videos, it is necessary to make sure that the image and video have been recorded by a camera, and that no artifact has been added. The detection of artifacts is a key element to use an image or a video in court. Thus, the integrity of images and videos used as a proof of evidence should be clearly assessed. The trustworthiness of images and videos has clearly an essential role in many security areas, including forensic investigation, criminal investigation, surveillance systems, and intelligence services.

As stated by Mahdian and Saic [21], verifying the integrity of digital images and detecting the traces of tampering without using any protecting pre-extracted or pre-embedded information have become an important research field of image processing. We will utilize and develop blind methods for detecting image forgery, that is, methods that use the image function to perform the forgery detection task. These methods are based on the fact that forgeries bring into the image-specific detectable changes (e.g., statistical changes). In high-quality forgeries, these changes cannot be found by visual inspection. Existing methods mostly try to identify various traces of tampering and detect them separately. The final decision about the forgery can be carried out by fusion of results of separate detectors.

Blind methods can be classified into several categories. In detection of near-duplicated image regions, a part of the image is copied and pasted into another part of the same image with the intention to hide an object or a region. There are methods capable of detecting near duplicated parts of the image that usually require a human interpretation of the results, see Refs. [21–23]. A different category includes interpolation and geometric transformation that are typically based on the resampling of a portion of an image onto a new sampling lattice, see, for example, Ref. [24]. In the photomontage detection problem, one of the fundamental tasks is the detection of image splicing, which can sometimes be based on analyzing the lighting conditions. Another category is related to compression method. In order to alter an image, typically the image is loaded to photoediting software, and once the changes are done, the digital image is resaved. Methods capable of finding the image compression history can be helpful in forgery detection. Another important category is the study of the noise characteristics and the chromatic aberrations [25, 26]. In the same line, blur and sharpening can also be analyzed to detect the concealment of traces of tampering. When two or more images are spliced together, it is often difficult to keep the appearance of the image correct perspective. Applying the principles from projective geometry to problems in image forgery detection can be also a proper way to detect traces of tampering. There are also other groups of forensic methods effective in forgery detection, see, for instance, single-view recaptured image detection, aliveness detection for face authentication, and device identification in digital image forensics, Refs. [27–30].

4.2 Case-based reasoning

4.2.1 State of the art

Case-Based Reasoning has been shown a successful problem-solving method in different applications were generalized knowledge is lacking. CBR has been used to interpret images [31, 32], 1-D signals [31, 33, 34], and text cases [35]. It also has been used for meta-learning of the best parameter of image segmentation [36] and classification methods [37], so that the best processing and classification results can be achieved, although domain knowledge is lacking. The success of these systems is because cases can be more easily collected than rules or other domain data and because of the flexibility of the systems based on their learning and maintenance mechanisms that allow incrementally improvement of their system performance during usage of the system.

4.2.2 Beyond the state of the art

The necessity to study the taxonomy of similarity measures and a first attempt to construct a taxonomy over similarity measures has been given by Perner [38] and has been further studied by Cunningham [39]. More work is necessary especially when not only one feature type and representation is used in a CBR system, as it is the case for multimedia data. These multimedia cases will be more complex as the cases used in the system described above that only face on one specific data type. To understand the similarity between these multimedia cases will require more complex knowledge of similarity by the police investigator for the different types of multimedia data. To develop novel similarity measures for text, videos, images, and audio and speech signals and to construct a taxonomy that allows understanding the relation between the different similarity measures will be a challenging task. Similarity aggregation of the different types of similarity measures is another challenging topic. Specific knowledge for the different types of data such as text [40, 41], images [42–44], video [45], 1-D signals, and meta-learning [36] is required in this work. The development of new similarity measures for multimedia data types and new data representations and ontologies will be done. A complex CBR system that can handle so many different data types, similarities, and data sources is a novelty.

Retrieval of multimedia data from a case base can be refined by relevance feedback mechanisms [46–52]. The user is asked to mark retrieved results as being "relevant" or not with respect to his/her interests. Then, feature weights and the similarity measures are suitably adapted to reflect user's interests. Relevance feedback can be implemented in a number of ways, for example, as the solution of an optimization problem, or as a classification problem. According to the problem at hand, the most suited formulation has to be devised. Thus, the main challenge will be to formulate the relevance feedback problem for forensic applications, so that the search is driven toward the cases more relevant to the case at hand.

Research has been described for learning of feature weights and similarity measures [53–55]. Case mining from raw data in order to get more generalized cases has been described by Jaenichen and Perner [56]. Learning of generalized cases and the hierarchy over the case base has been presented by the authors of Refs. [45, 57]. These works demonstrate that the system performance can be significantly improved by these functions of a CBR system.

New techniques for learning of feature weights and similarity measures and case generalization for different multimedia types are necessary and will be developed for these tasks.

The question of the Life Cycle of a CBR system goes along with the learning capabilities, case base organization and maintenance mechanism, standardization, and software engineering for which new concepts should be developed. As the result, we should come up with generic components for a CBR system for multimedia data analysis and interpretation that form a set of modules that can be easily integrated and updated into the CBR architecture. The CBR system architecture should easily allow configuring modules for new arising task.

The partner IBAI has a number of national and international patents that protect their work on CBR for images and signals. It is to expect that new methods will be developed that can be protected by patents and can ensure the international competition of European entities on CBR systems.

4.3 Multimedia feature extraction

4.3.1 State of the art

Most of computer vision algorithms rely on the extraction of meaningful features that transform raw data values into a more significant representation, better suited for classification and recognition. Although considered often not a central problem, the quality of feature representation can have critically important implications for the performance of the subsequent recognition methods.

Features are usually defined and selected according to a problem-oriented strategy, that is, ad hoc in light of the information considered relevant for the task at hand. In forensics, a plethora of features have been defined for the automated solutions to different problems, such as face detection, retrieval and recognition in video and images [58–60], individual people tracking over video sequences [61, 62], recognition of different biometric parameters (ear, gait, and iris) in images or videos [63, 64], speaker identification in audio signals, suspicious word detection, and handwriting recognition in text document.

Main challenges in forensics scenarios regard the unconstrained conditions in which multimedia data are collected. For audio signals, this is usually in the form of channel distortion and/or ambient noise. For videos and images, problems arise from changes in the illumination direction and/or in the pose of the subjects, occlusions, aging, and so on.

For images and videos, according to the problem at hand, the features selected can be based on specific morphologic parameters of individuals, such as face characteristics (e.g., nose width and eye distance) [65], posture and gesture, ear details, and so on or on general appearance features computed with low-level descriptors. These descriptors can be either global or local and can exhibit different degrees of invariance. Global descriptor category includes features based on Principal Component Analysis (PCA) [66] and Linear Discriminant Analysis (LDA) [67]. The local descriptor category is currently spreading and comprises features based on local values of color, intensity, or texture. To this category belong Scale-Invariant Feature Transform (SIFT) [68], Local Binary Pattern (LBP) [69], Histograms of Oriented Gradients (HOG) [70], or Gabor wavelets [71]. LBP is a well-known texture descriptor and a successful local descriptor robust to local illumination variations [72]. LBP descriptors are compact and easy to compare by various histogram metrics. In addition, there are many LBP variants that improve the description performance; among these, the most popular is Multi-Scale LBP (MSLBP) [73]. HOG has been successfully applied to tasks such as human detection [70] and face recognition [74]. Similar to LBP, edge information captured by gradients within blocks is packed into a histogram. Discarding pixel location information by block-based histogram binning, LBP and HOG gain invariance to local changes such as small facial expressions and pose variations in pedestrian images. The Gabor wavelets are also successful descriptors that capture global shape information centered at a pixel [75]. The convolution of multiple Gaussian-like kernels with different scales and orientations captures information insensitive to expression variation and blur at a pixel's location. Recently, a generalization of the Pairs of Pixels (POP) descriptor, called Centre Symmetric-Pairs of Pixels (CCS-POP), has been presented for face identification [76]. Another line of research currently gaining attention regards the computation of biologically inspired descriptors that result from the attempt to mimic natural visual systems. Several works have shown interesting results in a variety of different face and object recognition contexts [77–79].

The approach based on local descriptors has recently gained popularity, especially in relation to the spreading of the bag-of-feature representation. Indeed, in this frame, local feature descriptors, which can achieve high robustness with respect to appearance variations, are employed to develop a bag of descriptors that represent image content. All such descriptors are, then, quantized using learned visual words to facilitate the retrieval or classification [80–83]. The approach seems promising in forensic scenarios to fit the high variation of object appearance across different views since some very informative local features can accommodate to bad localizations or part visibility [62].

4.3.2 Beyond the state of the art

The problem of automatically extracting relevant information out of the enormous and steadily growing amount of electronic text data is becoming much more pressing. To overcome this problem, various technologies for information management systems have been explored within the Natural Language Processing (NLP) community. Two promising lines of research are represented by the investigation and development of technologies for (a) ontology learning from document collections and (b) feature extraction from texts.

Ontology learning is concerned with knowledge acquisition from texts as a basis for the construction of ontologies, that is, an explicit and formal specification of the concepts of a given domain and of the relations holding between them; the learning process is typically carried out by combining NLP technologies with machine learning techniques. Buitelaar [84] organized the knowledge acquisition process into a "layer cake" of increasingly complex subtasks, ranging from terminology extraction and synonym acquisition to the bootstrapping of concepts and of the relations linking them. Term extraction is a prerequisite for all aspects of ontology learning from text: measures for termhood assessment range from raw frequency to Information Retrieval measures such as TF-IDF, up to more sophisticated measures [85–88]. The dynamic acquisition of synonyms from texts is typically carried out through clustering techniques and lexical association measures [89, 90]. The most challenging research area in this domain is represented by the identification and extraction of relationships between concepts (taxonomical ones but not only); this research area presents strong connections with the extraction of relational information from texts, both relations and events (see below).

With feature extraction, we refer to the task of automatically identifying in texts instances of semantic classes defined in an ontology. This task includes recognition and semantic classification of items representing the domain referential entities ("Named Entity Recognition" or NER), either "named entities" or any kind of word or expression that refers to a domain-specific entity. Recently, extraction of inter-entity relational information is becoming a crucial task: relations to be extracted range from "place_of", "author_of," etc. to specific events, where entities take part in with usually predefined roles ("Relation Extraction"). Currently, there exist several feature extraction approaches, addressing different requirements, operating in different domains and on different text types, and extracting different information bits. If we look at the type of the underlying extraction methodology, systems can be classified into the following classes:

- rule-based systems, using hand-crafted rules. Rule-based systems are particularly appropriate for dealing with documents showing very regular patterns, such as standard tables of data, Web pages with HTML markup, or highly structured text documents;

- systems incorporating supervised machine learning: an alternative to the time-consuming process of hand coding of detailed and specific rules is represented by supervised semantic annotation systems, which learn feature extraction rules from a collection of previously annotated documents; and

- systems using unsupervised machine learning: they represent a viable alternative, currently being explored in different systems, to supervised machine learning approaches, as they dispense with the need for training data whose production may be as time consuming as rule hand coding.

Depending on nature and depth of the features to be extracted, different amounts of linguistic knowledge must be resorted to. This means that type and role of the linguistic analysis differ from one system to another. The condition part of feature extraction rules may check the presence of a given lexical item, the syntactic category of words in context, and their syntactic dependencies. Different clues such as typographical features, relative position of words, or even coreference relations can also be exploited. Most feature extraction systems therefore involve linguistic text processing and semantic knowledge: segmentation into words, morphosyntactic tagging, (either shallow or full) syntactic analysis, and sometimes even lexical disambiguation, semantic tagging, or anaphora resolution.

Text analysis can be carried out either at the preprocessing stage or as part of the feature extraction process. In the former case, the whole text is first analyzed. The analysis is global in the sense that items that are spread all over the document can contribute to build the normalized and enriched representation of the text. Then, the feature extraction process operates on the enriched representation of the text. In the latter case, text analysis is driven by the process of verifying a specific condition. The linguistic analysis is local, focuses on the context of the triggering item associated with a specific feature, and fully depends on the conditions to be checked for that feature.

Different approaches to feature extraction will be investigated to assess their strength and effectiveness to detect and describe the multimedia data content relevant to forensic activities. Both biometric features and local informative descriptors will be studied and collected to create a range of different opportunities to describe multimedia data content. More precisely, low level, local, invariant descriptors will be explored to assure a good performance of detection algorithms, especially for recognition in the wild, whereas global biometric features and properties will be considered as high-level information that is better understandable by end users.

A formal model will be adopted to define the features of different kinds. This will result into an ontological model that will organize different classes of features and foster their sharing and reuse. This will be a very innovative result since the ontology will be general and will approach the domain of multimedia data analysis. It will go further current metadata standards such as MPEG 7 or 21 and will be much more comprehensive and specific than other existing ontologies, which are only partially focused on feature extraction and always aimed at other problems such as multimedia data annotation. Additionally, the ontology will be enriched with algorithms to compute the features included, resulting into a toolbox for feature extraction. This will be another very innovative result.

As far as feature extraction from texts is concerned, the main challenge is represented by the typology of texts to be dealt with, testifying noncanonical language usages.

4.4 Text mining

4.4.1 State of the art

Twitter is a new multimedia communication channel that is rapidly gaining popularity and users, yet police forces do not dispose of adequate methods to analyze the large amounts of textual data that are generated each day. Recently, several retrospective investigations concerning football riots revealed that Twitter was actively used by rivaling gang members to plan their assaults. Twitter data are hard to analyze because the text fragments are very short, multiple persons can be involved in a conversation about various topics, and the data are rapidly changing.

Twitter is a recently introduced microblogging and information sharing platform [91] with over 140 million users and 340 million tweets per day. In the past, several studies have been dedicated to analyzing twitter feeds, for example, in the field of opinion mining and sentiment analysis. For example, in Ref. [92], the authors analyzed the text content of daily Twitter feeds by two mood tracking tools: OpinionFinder, which measures positive versus negative mood, and Google-Profile of Mood States (GPOMS), which measures mood in terms of six dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). They cross-validated the resulting mood time series by comparing their ability to detect the public's response to the presidential election and thanksgiving day in 2008. Ratkiewicz et al. [93] used machine learning for analyzing politically motivated individuals and organizations that use multiple centrally controlled twitter accounts to create the appearance of widespread support for a candidate or opinion and to support the dissemination of political misinformation.

4.4.2 Beyond the state of the art

We propose to develop and use an integrated data visualization environment based on formal concept analysis, temporal concept analysis, temporal relational semantic systems, and self-organizing maps to identify suspicious tweets.

Formal concept analysis (FCA) is a mathematical technique that was introduced in 1982 by Rudolf Wille [94] and takes its roots in earlier work of Birkhoff [95] and the early work on applying lattice-theoretical ideas in information science, like it was done by Barbut et al. [96]. FCA was used in several security text mining projects. The goal in each of these papers was to make an overload of information available in an intuitive visual format that may speed up and improve decision making by police investigators on where and when to act. In the first case study, with the Amsterdam-Amstelland police (RPAA), which started in 2007, FCA was used to analyze statements made by victims to the police. The concept of domestic violence was iteratively enriched and refined, resulting in an improved definition and highly accurate automated labeling of new incoming cases [97]. Later on, the authors made a shift to the millions of observational and very short police reports from which persons involved in human trafficking and terrorism were extracted. Concept lattices allowed for the detection of several suspects involved in human trafficking or showing radicalizing behavior [98, 99].

Temporal concept analysis (TCA) was introduced by Wolff [100] and offers a framework for representing and analyzing data containing a temporal dimension. In previously discussed security applications, suspects were mentioned in multiple reports, and a detailed profile of one suspect (and persons in his social network) depicted as a lattice, with timestamps of the observations as objects and indications as attributes helped to gain an insight into his (their) threat to society [101]. Recently, TCA and its relational counterpart temporal relational semantic systems (TRSS, [100]) were successfully applied to the analysis of chat conversations [102].

Self-organizing maps [103] have been used in many applications, where high-dimensional unsupervised data spaces had to be visualized in a two-dimensional plane to make the data accessible for human experts. For example, Ramadas et al. [104] used self-organizing maps for identifying suspicious network activity. In a previous security case study, a special type named emergent self-organizing maps was used to identify domestic violence in police reports [105, 106]. They were found to be more suitable than multidimensional scaling for text mining. Claster et al. [107] used self-organizing maps to mine over 80 million twitter micro logs in order to explore whether these data can be used to identify sentiment about tourism and Thailand amid the unrest in that country during the early part of 2010 and further whether analysis of tweets can be used to discern the effect of that unrest on Phuket's tourism environment.

Nevertheless, there are several differences between analyzing twitter feeds and traditional police reports. Whereas individual tweets may not be so interesting, a lot of information can be distilled from conversations consisting of many tweets that emerged between different users concerning a certain topic. Such feeds do not contain a summary of facts; rather several topics emerge between two or more persons. We should judge the interestingness of the feed from a security enforcement perspective and distinguish between several types of twitter users in a relevant conversation, for example, is this person someone who contributed only marginally or did he or she actually contribute to or promote criminal behavior. Ebner et al. [108]

used Formal Concept Analysis (FCA) to categorize twitter users who write tweets about the same topics in the context of a conference event. Cuvelier et al. [109] used FCA as an e-reputation monitoring system in combination with tag clouds. Also, the Natural Language Processing of tweets is nowadays a challenging task since Twitter is characterized by a so-called noncanonical language. It is widely acknowledged that NLP systems have a drop of accuracy when tested against text characterized by this kind of language. This negatively affects different levels of text analysis ranging from the linguistic annotation to the information extraction process. It follows that the analysis of noncanonical languages is one of the main topics of the most recent NLP conferences, for example, the First Workshop on Syntactic Analysis of Noncanonical Language (SANCL-2012) (https://sites.google.com/site/sancl2012/), the workshop series on Scritture brevi (lit.: short writings) organized by the University of Rome Tor Vergata (https://sites.google.com/site/scritturebrevi/atti-dei-workshop), and the First Shared Task on Dependency Parsing of Legal Texts at SPLET-2012 (https://sites.google.com/site/splet2012workshop/shared-task). The main challenges in analyzing noncanonical languages, as tweet language, result from the fact that they have different linguistic characteristics with respect to the data from which the tools are trained, typically newswire texts. Among the others, punctuation and capitalization are often inconsistent; slang, technical jargon is widely exploited; and noncanonical syntactic structures frequently occur [110–112]. Accordingly, several domain adaptation methods and different strategies of analysis have been investigated to improve the accuracies of the NLP tools, among the most recent ones the self-training method used by Le Roux et al. [113], the active-learning method used by Attardi et al. [114], and the term-extraction method proposed by Bonin et al. [88].

Event detection in Twitter has been recently an area of active research and successfully applied to detecting earthquakes [115] and sport events [116]. For events of interest to legal forces, one can utilize the generic features, such as emerging common terms, location, date, and also potentially the participants of the event. Hence, we extract the date/time information and time-event phrases that are learnt from tweets and set the presence of them as a feature. Participant information is also captured via the presence of the '@' character followed by a username within tweets. Specific to the events of legal interest, one can also utilize the overall sentiment of the tweets as a potential feature. According to a recent research by Leetaru [117] at the University of Illinois at Chicago, strong negative emotions in news can suggest upcoming of a significant event. A sentiment analysis in a long period of news revealed that the textual sentiments before the revolutions in Libya and Egypt have shown significant negative signals. The strength of this negativity is found comparable to the signals in 1991 news, right before the United States entered Kuwait; and also in 2003, when the United States-Iraq was about to start.

While the current approaches, such as Ref. [117], have been shown to work on static data and static models, more research is needed to enhance these methods for the dynamic case. Also, the news text is highly structured and formal, while Twitter consists of informal short text. Based on our prior work on classifying short tweets [118], and sentiment analysis on large-scale data [119], we will categorize the tweets for event detection and identify tweets with strong sentimentality. Our initial hypothesis is that strong sentiment increases the probability of event being of interest to legal forces. Recently, distributional semantic models (DSMs) have been applied to affective text analysis with good results across languages [120]. In this WP, we will also apply DSMs to sentiment analysis of multilingual tweets. The more interesting problem is the forecasting problem, where the events can be predicted beforehand. This would be of high value for preventive law enforcement. Besides the prediction problem, one can also use this approach to get feedback from the crowd on actions taken by the law officers. Such approaches have already been deployed for finance and marketing applications to understand the mood of financial markets and consumer opinions [92, 121, 122]. Similar concepts can be adapted for forensic applications. In fact, FBI and Pentagon have already started to utilize these methods to predict criminal and terrorist activities and monitor persons and regions of high interest [AP Exclusive].

The innovativeness of tool in this area lays in the fact that the combination of the discussed methods has never been proposed for visualizing and clustering data, nor integrated in a software system. It will be the first integrated human centered data discovery environment that combines both statistical methods from machine learning with order-theoretic methods such as concept lattices. The self-organizing map that can handle high-dimensional data spaces and, as a consequence, is an ideal tool for an initial preprocessing is at the start of the human centered discovery process. FCA can then be used to explore dependencies and information links in a smaller subset. TCA and TRSS are used for in-depth profiling of identified individuals and communities. In particular, we focus on the niche of twitter user and feed mining in the broader text-mining field. State-of-the-art domain adaptation methods will be tested to improve the accuracies of the linguistic annotation tools on Twitter data, and customized term-extraction methods will be devised in order to reliably extract relevant keywords from tweets. Needless to say that the proposed system can be easily expanded to other text mining applications.

A web crawler will be designed to collect the feeds from the twitter website. This is a technically challenging yet known task to the scientific community (see e.g., [107]). The data collection can be done by an employee hired by the police who received a type P screening. The type of data is fragments of texts. Concerning languages, we will first focus on Dutch tweets. This may later be extended to Hungarian and Bulgarian since most organized crime in areas such as human trafficking is committed by these nationalities in Amsterdam. Since a tweet consists of among others a user name, his twitter ID and the posted text, as well as potentially ID and name of other users, we will first replace these user-identifiable information items by numeric values using regular expressions. In the second step, we will use available Named Entity Recognition methods for removing person names from the tweets themselves.

4.5 Video analysis

4.5.1 State of the art

Video retrieval has a long history [123–125]. According to the type of video at hand (e.g., film, news, CCTV recording, etc.), different retrieval tasks can be defined both in terms of the type of query and in terms of the processing techniques that are suited for extracting meaningful concepts. For example, it is easy to see that the making of a film comprises the use of techniques whose goal is to provoke sentiments in the watcher. Thus, in order to retrieve concepts from videos, automatic techniques must take into account not only the characteristic of the scene but also the movements of the camera and video editing techniques. On the other hand, still cameras used for video-surveillance purposes allow for the detection of persons and objects moving within the monitored area, as the characteristic of the scene is well known in advance. On these topics, a vast corpus of research has been carried out in the past years, and a number of automatic analysis techniques are embedded into commercial products [126].

One of the first steps in video analysis is the detection of shots, that is, video sequences that contain a continuous camera action in time and space [127, 128].

In the case of films, broadcasted news, and sport videos, shot detection is performed by looking at well-known separators, such as fading and black frames. Each shot is then characterized by one or more key frames, that is, those frames that can be used to characterize the shot. Shot classification can be performed by extracting suitable features and using machine-learning techniques for concept classification. Features can be either extracted from key frames, as well as by looking at global characteristics of the video sequence. They can represent low-level information of such as color and textures as well as characteristics of the shot such as temporal features.

A number of techniques for carrying out these steps have been developed for TV broadcasters, in particular for sport as well as news programs [123, 124]. In these areas, the knowledge of the rules of the game and the rules of video shooting allowed for building a reliable ground truth that allows to make objective comparisons of different algorithms. The classification of video shots can be used for retrieval purposes, as soon as the goal is to retrieve all videos related to a particular class. On the other hand, the use of these techniques for forensic applications still needs more investigation due to the low resolution of the cameras, the variability of the recorded scenes, and the presence of person and objects typically in nonfrontal positions and with many occlusions.

Today, it is of particular interest the reidentification of people in videos [129, 130]. This problem can be formulated as follows. In many real scenarios, an area is monitored by a number of cameras. When persons move in the monitored environment, they can be identified by their face only if they appear in the video in some pose. After they have been identified in one of the videos, they can be tracked (i.e., reidentified) according to their global appearance (e.g., their clothes) rather than by their face.

Speech and sound files constitute an important part of the data collected by Law Enforcement Agencies. For the last 35 years, practical speech recognition systems have been based on Hidden Markov Models (HMMs), which model the training data using the Baum-Welch algorithm in a global manner. Markov state probability distributions are also represented using Gaussian Mixture Models (GMMs). HMMs try to represent the time-varying speech and sound files [131, 132]. This approach is successful to some extent in controlled environments and dictation systems in which people clearly speak to the machines [133].

HMMs and GMMs use features extracted from temporal speech windows. Current speech and sound feature extraction schemes are based on Fourier analysis [131, 134, 135]. Temporal information is only incorporated to the automatic speech recognition systems by only dividing speech into temporal analysis windows. Unfortunately, this global approach loses keyword or speaker-specific features, which are needed in forensic applications. For example, a person cannot modify his or her own average temporal zero crossing rate, even if he or she tries to change his or her own voice by mumbling, or talking with a mouth full of food or cotton balls, etc. [136]. This kind of temporal and person specific information is not used in today's systems, which are globally trained using all the available data.

Global approaches provide good speech and speaker recognition and identification results as long as it is possible to have a good description of the unobserved data. However, continuous spontaneous speech recognition is still an unsolved problem [133, 137]. Unfortunately, most of the speech data in legal cases are spontaneous speech data. In many applications, it is required to retrieve keywords, phrases, names, and speakers from spontaneous speech in real time. Therefore, it is necessary to develop not only new feature extraction and speech and sound representation schemes but also exemplar type case and similarity-based reasoning methods to improve the current speech and sound processing systems.

4.5.2 Beyond the state of the art

The analysis of videos for forensic applications can be carried out by relying on some of the above techniques, provided they are tailored to the scenario at hand. It is easy to see that in the case of surveillance videos, we cannot define a shot according to the paradigm used to segment a film or a sport video [126, 138]. Rather, the definition of "shot" can be driven by the event that is looked for in the video. In particular, the video analyst should be able to query the system, so that the video is first segmented according to the particular event, and then, the shots that can contain the event of interest with high probability are further analyzed by more sophisticated technique in order to detect the object of interest [139]. The development of such a system is beyond the current state of the art, and it will be carried out within this project.

The development of reidentification techniques may allow tracking a person in videos collected by multiple cameras at different locations and in different periods. Detecting people can be carried out by face detection. Many of the existing facial recognition systems are sensitive to variations in the enrolment phase [140–145]. Often these systems have been trained by a huge number of pictures of the same person to estimate reliable values of the parameters for statistical classifiers. The current state of the art does not include a suitable system for the generation of a prototype picture of a person nor a suited prototype-based classifier [146, 147]. Some automatic prototype generation developed in the area of pattern recognition could be used for face recognition [148–150].

Prototype-based system could effectively handle changes in illumination, as they can perform recognition by part resemblance [151, 152]. For the above reasons, most of the facial recognition systems available today assume a standardized enrolment procedure to be performed in a controlled environment (e.g., a cabin), where a number of pictures of the face in a frontal position (2-D) with respect to the camera are taken. In addition, the picture is renewed whenever the recognition accuracy decreases.

Many different methods have been used so far for face recognition and cover a wide spectrum of methods in the pattern recognition field: geometrical representation of the face [153], templates [154, 155], hidden Markov models [156], principal component analysis [157], independently component analysis [158], elastic graph matching [159], trace transform [160], and SVM [161]. None of the methods can be seen as the most promising method because the performance depends on the scenario at hand, and the assumption behind the proposed theoretical models might not be met in real scenarios. Thus, new techniques based on the exploitation of different picture representations, such as shape, texture, signs for skin, eyes and spatial, sign-based connections, and the prototype-based system, have to be investigated.

Case and similarity-based recognition and sensing methods for speech, sound, and audio recognition using both temporal and frequency domain information will be developed. Development of "query by example," keyword, and phrase-based retrieval schemes using exemplar-based schemes, which will be capable of part and whole similarity matching, will be a significant contribution to the existing speech recognition systems.

Current methods for speech and audio analysis emphasize spectral methods. For example, well-known Shazam music recognition method uses only spectral peaks [162]. Commonly used mel-cepstral coefficients, line spectral frequencies, and RASTA features [134, 135] do not have any temporal information, either. We believe that temporal information is not fully utilized in current methods. Temporal information will provide critical information for speaker recognition and keyword spotting applications. We are developing temporal speech representation methods based on delta modulation [163, 164], zero-crossing, and wavelet scattering [165, 166] information will be incorporated into content based audio and sound retrieval and speech and audio recognition applications.

As pointed above, another important avenue, which is not explored by current methods, is compressive recognition, similarity-based reasoning, and case-based reasoning. Current data modeling methods assume a global representation. On the other hand, case and similarity-based reasoning methods will be able to incorporate fine details of the test case and will likely to provide better recognition results, especially in spontaneous speech. Temporal representation methods such as delta modulation and zero-crossing information are ideal for exemplar and similarity-based reasoning approaches. It is also possible to combine the differential representation of temporal data with the spectral data using compressive sensing [167], which extends this differential data processing concept by using random weights adding to zero to linearly combine the data and/or features. In this way, similarity learning, case generalization and case storage, and compressive learning and sensing will allow the handling of very large amount (terabytes) of data. Once the keyword and phrases are detected, analysts can manually process the proposed retrieval results.

Cut-and-paste locations in speech can be also detected using delta modulation and wavelet scattering, providing a differential representation of speech, sound, and audio data. Fragile watermarking schemes based on wavelet scattering and delta modulation will be developed to prevent tampering. Resulting representation can be easily stored, and it will be ideal for different forensic purposes.

5. Conclusions

Forensic investigations on multimedia evidence usually develop along four different steps: analysis, selection, evaluation, and comparison. During the analysis step, technicians typically look at huge amounts of different multimedia data (e.g., hours of video or audio recordings, pages and pages of text, and hundreds and hundreds of pictures) to reconstruct the dynamic of the event and collect any piece of relevant information. This step obviously requires a lot of time, and many factors can make it difficult, among which data heterogeneity, quality, and quantity are the most relevant. Afterward, during the selection step, technicians select and acquire the most meaningful pieces of information from the different multimedia data (e.g., frames from videos, audio fragments, and documents). Then, in the evaluation step, they look for relevant elements in the selected data, which will be further investigated in the comparison step. They can select heads, vehicles, license plates, guns, sentences, sounds, and all other elements that can link a person to the event. The main problems are the low quality of media data due to high compression, adverse environmental conditions (e.g., noise, bad lighting condition), camera/object position, and facial expressions. Finally, during the comparison step, technicians place the extracted elements side by side with a known element of comparison. From the comparison of general and particular characteristics, the operators give a level of similarity. In forensic application, the use of automatic pattern recognition system gives poor performance because of the high variability of data recording. On the other hand, human perception is a great pattern recognition system but is characterized by high subjectivity and unknown reproducibility and performance.

In this chapter, we propose to develop a toolkit of methods and instruments that will be able to support analysts along all these steps, strongly reducing human intervention. First of all, it will include instruments to process different kinds of media data and, possibly, correlate them. This will obviously reduce the time spent to find the correct instruments for processing the medium at hand. Furthermore, it comprises preprocessing tools that alleviate, by filtering and enhancement, the problem of low-data quality. In particular, for image and video data, a great help will come from super-resolution methods that will maximize the information contained in low-resolution images or videos (e.g., foster the process of face reconstruction and recognition from blurred images). This feature will greatly support all the subsequent steps.

In this chapter, we focused on the background and motivation for our work. The overall system architecture is explained. We present the data to be used. After a review of the state of the art of related work of the multimedia data we consider in this work, we describe the method and techniques we are developing that go beyond the state of the art. The work will be continued in the Chapter Part II of Forensic Multimedia Data Analysis.

Author details

Petra Perner
Institute of Computer Vision and Applied Computer Sciences, Leipzig, Germany

*Address all correspondence to: pperner@ibai-institut.de

References

[1] Passive millimeter wave images, copyright. Alfa Imaging S.A. ALFA, Spain

[2] Molina R, Murtagh F, editors. DSP soars into space. IEEE Signal Processing Magazine. 2001;**18**(2):1-3

[3] Babacan SD, Molina R, Do MN, Katsaggelos AK. Blind deconvolution with general sparse image priors. In: European Conference on Computer Vision (ECCV), Berlin, Heidelberg, Florence, Italy: Springer; 2011. pp. 984-999

[4] Starck JL, Murtagh F, Candès EJ, Donoho DL. Gray and color image contrast enhancement by the curvelet transform. IEEE Transactions on Image Processing. 2003;**12**:706-717

[5] Daubos T, Murtagh F. High-quality still images from video frame sequences. In: Geradts Z, Rudin LI, editors. Investigative Image Processing II, Proceedings of the SPIE. Vol. 4709. 2002. pp. 49-59

[6] Daubos T, Geradts Z, Starck JL, Campbell J, Murtagh F. Automated wavelet-based image addition: application to surveillance video. In: Whelan PF, editor. IMVIP'99—Irish Machine Vision and Image Processing Conference 1999, Dublin City University. 1999. pp. 15-25

[7] Luessi M, Babacan SD, Molina R, Booth JR, Katsaggelos AK. Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors. NeuroImage. 2011;**55**(1):113-132

[8] Vega M, Mateos J, Molina R, Katsaggelos AK. Bayesian TV Denoising of SAR images. In: IEEE International Conference on Image Processing ICIP 2011. Bruselas (Bélgica); 2011. pp. 169-172

[9] Amizic B, Spinoulas L, Molina R, Katsaggelos AK. Compressive sampling with unknown blurring function: Application to passive millimiterwave imaging. In: IEEE International Conference on Image Processing. Orlando, Florida; 2007. pp. 321-329

[10] Babacan SD, Molina R, Katsaggelos AK. Total variation image restoration and parameter estimation using variational posterior distribution approximation. In: International Conference on Image Processing (ICIP)—IBM Student Paper Award for ICIP 2007. Vol. I. San Antonio, Texas (USA); 2007. pp. 97-100

[11] Starck J-L, Murtagh F, Fadili J. Sparse Image and Signal Processing: Wavelets. Curvelets: Morphological Diversity. Cambridge University Press; 2010

[12] Starck J-L, Murtagh F. Astronomical Image and Data Analysis. 2nd ed. Springer-Verlag; 2006

[13] Chantas G, Galatsanos N, Molina R, Katsaggelos AK. Variational Bayesian image restoration with a spatially adaptive product of total variation image priors. IEEE Transactions on Image Processing. 2010;**19**(2):351-362

[14] Babacan D, Molina R, Katsaggelos AK. Bayesian blind deconvolution from differently exposed image pairs. IEEE Transactions on Image Processing. 2010;**19**(11):2874-2888

[15] Katsaggelos AK, Molina R, Mateos J. Super resolution of images and video. In: Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool; 2007

[16] Babacan SD, Molina R, Katsaggelos AK. Variational Bayesian super resolution. IEEE

Transactions on Image Processing. 2011;**20**(4):984-999

[17] Mateos J, Katsaggelos AK, Molina R. A Bayesian approach to estimate and transmit regularization parameters for reducing blocking artifacts. IEEE Transactions on Image Processing. 2000;**9**(7):1200-1215

[18] Bishop C. Pattern Recognition and Machine Learning. Springer; 2006

[19] Barber D. Bayesian Reasoning and Machine Learning. Cambridge University Press; 2012

[20] Murphy KP. Machine Learning: A Probabilistic Perspective. MIT; 2012

[21] Mahdian B, Saic S. A bibliography on blind methods for identifying image forgery. Signal Processing: Image Communication. 2010;**25**(6):389-399

[22] Poisel R, Tjoa S. Forensics investigations of multimedia data: A review of the state-of-the-art. In: Proc. Sixth Int. IT Security Incident Management and IT Forensics (IMF) Conf. 2011. pp. 48-61

[23] Farid H. Exposing digital forgeries in scientific images. In: ACM MM&Sec. 2006

[24] Popescu AC, Farid H. Exposing digital forgeries by detecting traces of resampling. IEEE Transactions on Signal Processing. 2005;**53**(2):758-767

[25] Chen M et al. Determining image origin and integrity using sensor noise. IEEE Transactions on Information Forensics and Security. 2008;**3**(1):74-90

[26] Li C-T, Satta R. An empirical investigation into the correlation between vignetting effect and the quality of sensor pattern noise. IET Computer Vision. 2012;**6**(6):560-566

[27] Gao X, Ng TT, Qiu B, Chang S-F. Single-view recaptured image detection based on physics-based features. In: IEEE International Conference on Multimedia & Expo (ICME). 2010

[28] Ng T-T. Camera response function signature for digital forensics Part II: Signature extraction. In: IEEE Workshop on Information Forensics and Security (WIFS). 2009

[29] Khanna N, Delp EJ. Source scanner identification scanned documents. In: IEEE Workshop on Information Forensics and Security (WIFS). 2009

[30] Khanna N et al. Forensic techniques for classifying scanner, computer generated and digital camera images. In: IEEE ICASSP. 2008

[31] Perner P, editor. Case-Based Reasoning for Image and Signals, Series Computational Intellignece. Berlin: Springer Verlag; 2007

[32] Perner P, Holt A, Richter M. Image processing in case-based reasoning. The Knowledge Engineering Review. 2005;**20**(3):311-314

[33] Ahmed MU, Begum S, Funk P. An overview of three medical applications using hybrid case-based reasoning. In: Perner P, editor. ICDM 2012, Workshop Proceedings, Workshop on Case-Based Reasoning CBR-MD 2012. Fockendorf: IBAI-Publishing; 2012. pp. 79-94 ISBN 978-3-942952-16-3

[34] Perner P, Attig A, Machno O. Novel method for the interpretation of spectrometer signals based on delta-modulation and similarity determination. Transactions on Mass-Data Analysis of Images and Signals. 2011;**3**(1):3-14

[35] Weber RO, Ashley KD, Breueninghaus S. Textual case-based reasoning. The Knowledge Engineering Review. 2006;**20**(3):255-260

[36] Attig A, Perner P. Model building in image processing by meta-learning based on case-based reasoning. In: Wang PS-P, editor. Pattern Recognition and Machine Vision-In Honor and Memory of Late Prof. King-Sun Fu, River Publishers' Series in Information Science and Technology. River Publishers; 2010. pp. 149-164

[37] Murtagh F, Starck JL. Wavelet and curvelet moments for image classification: Application to aggregate mixture grading. Pattern Recognition Letters. 2008;**29**:1557-1564

[38] Perner P. Why case-based reasoning is attractive for image interpretation. In: Perner P, Aha D, Watson I, editors. Case-Bases Reasoning Research and Developments, LNAI. Vol. 2080. Heidelberg: Springer; 2001. pp. 27-44 (invited paper)

[39] Cunningham P. A taxonomy of similarity mechanisms for case-based reasoning. IEEE Transactions on Knowledge and Data Engineering. 2009;**21**(11):1532-1543

[40] Iosif E, Potamianos A. Similarity computation using semantic networks created from web-harvested data. Natural Language Engineering. 2015:49-79

[41] Iosif E, Potamianos A. Unsupervised semantic similarity computation between terms using web documents. IEEE Transactions on Knowledge and Data Engineering. 2010;**22**(11):1637-1647

[42] Perner P, Perner H, Jänichen S. Recognition of airborne fungi spores in digital microscopic images. Journal Artificial Intelligence in Medicine, Special Issue on CBR. 2006;**36**(2):137-157

[43] Geradts Z, Bijhold J, Hermsen R, Murtagh F. Image matching algorithms for breech marks and firing pins in a database of spent cartridge cases of firearms. Forensic Science International. 2001;**119**:97-106

[44] Geradts Z, Bijhold J, Hermsen R, Murtagh F. Matching algorithms using wavelet transforms for a database of spent cartridge cases of firearms. In: Proceedings of SPIE. Vol. 4232. 2001. pp. 545-552

[45] Contreras P, Murtagh F. Fast, linear time hierarchical clustering using the Baire metric. Journal of Classification. 2012;**29**:118-143

[46] Thomee B, Lew M. Interactive search in image retrieval: a survey. Journal of Multimedia Information Retrieval. 2012;**1**(2):71-86

[47] Piras L, Giacinto G, Paredes R. Enhancing image retrieval by an exploration-exploitation approach. In: Perner P, editor. Machine Learning and Data Mining in Pattern Recognition, LNCS. Vol. 7376. Berlin: Springer; 2012. pp. 355-365

[48] Datta R, Joshi D, Li J, Wang JZ. Image retrieval: ideas, influences, and trends of the new age. ACM Computing Surveys. 2008;**40**:1-60

[49] Giacinto G, Roli F. Instance-based relevance feedback in image retrieval using dissimilarity spaces. In: Perner P, editor. Case-Based Reasoning for Signals and Images. Berlin: Springer-Verlag; 2007. pp. 419-430

[50] Giacinto G. A nearest-neighbor approach to relevance feedback in content based image retrieval. In: Proceedings of the 6th ACM International conference on Image and video retrieval (CIVR'07). ACM Press; 2007. pp. 456-463

[51] Tronci R, Murgia G, Pili M, Piras L, Giacinto G. ImageHunter: A novel tool for relevance feedback in content based image retrieval. In: Loi C, Semeraro G, Vargiu E, editors. New Challenges in

Distributed Information Filtering and Retrieval, SCI. Vol. 439. Heidelberg: Springer; 2013. pp. 53-70

[52] Lew MS, Sebe N, Djeraba C, Jain R. Content-based multimedia information retrieval: state of the art and challenges. ACM Transactions on Multimedia Computing, Communications, and Applications. 2006;**2**:1-19

[53] Craw S. Introspective learning to build case-based reasoning (CBR) knowledge containers. In: Perner P, Rosenfeld A, editors. Machine Learning and Data Mining in Pattern Recognition, LNCS. Vol. 2734. Heidelberg: Springer; 2003. pp. 1-6

[54] Wettschereck D, Aha DW, Mohri T. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artificial Intelligence Review. 1997;**11**:273-314

[55] Zhang L, Coenen F, Leng P. Formalising optimal feature weight settings in case-based diagnosis as linear programming problems. Knowledge-Based Systems. 2002;**15**:391-298

[56] Jaenichen S, Perner P. Conceptual clustering and case generalization of two-dimensional forms. Computational Intelligence. 2006;**22**(3/4):178-193

[57] Perner P. Case-base maintenance by conceptual clustering of graphs. Engineering Applications of Artificial Intelligence. 2006;**19**(4):381-393

[58] Schwartz W, Guo H, Choi J, Davis L. Face identification using large feature sets. IEEE Transactions on Image Processing (TIP). 2012;**21**(4):2245-2255

[59] Tolba AS, El-baz AH, El-Harby AA. Face Recognition: A literature review. International Journal of Signal Processing. 2006;**2**(2):88-103

[60] Jain AK, Klare B, Park U. Face matching and retrieval in forensics applications. IEEE MultiMedia. 2012;**19**(1):20-28

[61] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: Proc. ECCV 2008. 2008. pp. 262-275

[62] Liu K, Yang J. Recognition of people reoccurrences using bag-of-features representation and support vector machine. In: Chinese Conference on Pattern Recognition, Nanjing, 2009. pp. 1-5. DOI: 10.1109/CCPR.2009.5344034

[63] Sanderson C. Biometric Person Recognition: Face, Speech and Fusion. VDM Verlag; 2008

[64] Ali H, Salami MJE. Wahyudi: Iris recognition system by using support vector machines. In: International Conference on Computer and Communication Engineering, ICCCE; 2008. pp. 516-521

[65] Vezzetti E, Marcolin F. 3D human face description: Landmarks measures and geometrical features. Image and Vision Computing. 2012;**30**:698-712

[66] Turk M, Pentland A. Face recognition using eigenfaces. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. 1991. pp. 586-591

[67] Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997;**19**(1):11-720

[68] Lowe DG. Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision. Vol. 2. 1999. pp. 1150-1157

[69] Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;**28**(12):2037-2041

[70] Dalal N, Triggs B. Histograms of oriented gradients for human detection. CVPR. 2005

[71] Shen L, Bai L. A review on Gabor wavelets for face recognition. Pattern Analysis and Applications. 2006;**9**:273-292

[72] Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns. Pattern Recognition. 2009;**42**:425-436

[73] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;**24**(7):971-987

[74] Schwartz WR, Guo H, Davis LS. A robust and scalable approach to face identification. ECCV. 2010

[75] Zhu ZF, Tang M, Lu HQ. A new robust circular Gabor based object matching by using weighted Hausdorff distance. Pattern Recognition Letters. 2004;**25**(4):515-523

[76] Choi J, Schwartz WR, Guo H, Davis LS. A complementary local feature descriptor for face identification. In: IEEE Workshop on the Applications of Computer Vision (WACV). 2012

[77] Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007;**29**(3):411-426

[78] Mutch J, Lowe DG. Object class recognition and localization using sparse features with limited receptive fields. IJCV. 2008

[79] Cox D, Pinto N. Beyond simple features: A large-scale feature search approach to unconstrained face recognition. In: IEEE Int. Conference on Automatic Face & Gesture Recognition. 2011. pp. 8-15

[80] Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos. In: Proc. ICCV 2003. Nice, France; 2003. pp. 11-17

[81] Moosmann F, Nowak E, Jurie F. Randomized clustering forests for image classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008;**30**(9):1632-1646

[82] Colantonio S, Martinelli M, Salvetti O. Ontology and algorithms integration for image analysis. In: Salerno E, Çetin AE, Salvetti O, editors. Muscle 2011, LNCS. Vol. 7252. Berlin Heidelberg: Springer-Verlag; 2012. pp. 17-29

[83] Perner P. Image mining: Issues, framework, a generic tool and its application to medical-image diagnosis. Journal Engineering Applications of Artificial Intelligence. 2002;**15**(2):105-216

[84] Buitelaar P, Cimiano P, Magnini B, editors. Ontology Learning from Text: Methods, Evaluation and Applications Frontiers in Artificial Intelligence and Applications Series. Vol. 123. IOS Press; 2005

[85] Frantzi K, Ananiadou S. The C–value/NC value domain independent method for multi-word term extraction. Journal of Natural Language Processing. 1999;**6**(3):145-179

[86] Dell'Orletta F, Lenci A, Marchi S, Montemagni S, Pirrelli V, Venturi G. Dal

testo alla conoscenza e ritorno: estrazione terminologica e annotazione semantica di basi documentali di dominio. In: AIDA Informazioni, Atti del Convegno Nazionale Ass.I.Term "I–TerAnDo", Università della Calabria, 5-7 giugno 2008. Roma: AIDA, n. 1-2/2008, ISSN 1121-0095; 2008. pp. 185-206

[87] Lenci A, Montemagni S, Pirrelli V, Venturi G. Ontology learning from Italian legal texts. In: Breuker J et al., editors. Law, Ontologies and the Semantic Web—Channelling the Legal Information Flood, Frontiers in Artificial Intelligence and Applications. Vol. 188. Heidelberg: Springer; 2009. pp. 75-94

[88] Bonin F, Dell'Orletta F, Venturi G, Montemagni S. A contrastive approach to multi–word extraction from domain–specific corpora. In: Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010). La Valletta, Malta; 2010

[89] Lin D. Automatic retrieval and clustering of similar words. In: Proceedings of COLING/ACL98. Montreal, Canada; 1998

[90] Allegrini P, Montemagni S, Pirrelli V. Example-based automatic induction of semantic classes through entropic scores. In: Linguistica Computazionale. Vol. XVI–XVII. 2003. pp. 1-45

[91] Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on World wide web (WWW '10). New York, NY, USA: ACM; 2010. pp. 591-600

[92] Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. Journal of Computational Science. 2011;**2**(1):1-8

[93] Ratkiewicz J et al. Detecting and tracking political abuse in social media. In: Proc. of ICWSM. 2011

[94] Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival I, editor. Ordered Sets. Dordrecht, Boston: Reidel; 1982. pp. 445-470

[95] Birkhoff G. Lattice Theory. 3rd ed. Vol. 25. Providence, RI: American Mathematical Society Coll. Publ; 1973

[96] Kimberly Dozier AP exclusive: CIA following Twitter, Facebook. Available from: http://www.guardian.co.uk/world/feedarticle/9929898 [Accessed: 08 October 2012]

[97] Poelmans J, Elzinga P, Viaene S, Dedene G. Formally analyzing the concepts of domestic violence. Expert Systems with Applications. 2011;**38**(4):3116-3130. DOI: 10.1016/j.eswa.2010.08.103

[98] Elzinga P, Poelmans J, Viaene S, Dedene G, Morsing S. Terrorist threat assessment with formal concept analysis. In: Proc. 8th IEEE International Conference on Intelligence and Security Informatics. 23-26 May. Vancouver, Canada; 2010. pp. 77-82 ISBN: 978-1-42446460-9/10

[99] Poelmans J, Elzinga P, Viaene S, Dedene G, Kuznetsov S. Semi-automated knowledge discovery in unstructured text: Identifying and profiling human trafficking. International Journal of General Systems. 2012;**41**(8):774-804

[100] Wolff KE. Temporal concept analysis. In: Nguifo EM et al, editors. ICCS-2001 International Workshop on Concept Lattices-Based Theory, Methods and Tools for Knowledge Discovery in Databases, Stanford University. Palo Alto, CA; 2001. pp. 91-107

[101] Poelmans J, Elzinga P, Viaene S, Dedene G, Kuznetsov S. A concept discovery approach for fighting human

trafficking and forced prostitution. In: 19th International Conference on Conceptual Structures, July 25-29, Derby, England, LNCS. Vol. 6828. Heidelberg: Springer; 2011. pp. 201-214

[102] Elzinga P, Wolff KE, Poelmans J. Analyzing chat conversations of pedophiles with temporal relational semantic systems. In: 1st IEEE European Conference on Intelligence and Security Informatics. Odense, Denmark; 22-24 August 2012. 2012. pp. 242-249

[103] Kohonen T. Self-organized formation of topologically correct feature maps. Biological Cybernetics. 1982;**43**:59-69

[104] Ramadas M, Ostermann S, Tjaden B, Vigna G, Kruegel C, Jonsson E. Detecting anomalous network traffic with self-organizing maps. In: Recent Advances in Intrusion Detection, LNCS. Vol. 2820. Heidelberg: Springer; 2003. pp. 36-54

[105] Poelmans J, Elzinga P, Viaene S, Van Hulle M, Dedene G. Text mining with emergent self organizing maps and multi-dimensional scaling: A comparative study on domestic violence. Applied Soft Computing. 2011;**11**(4):3870-3876. DOI: 10.1016/j.asoc.2011.02.026

[106] Poelmans J, Elzinga P, Viaene S, Van Hulle M, Dedene G. Gaining insight in domestic violence with emergent self-organizing maps. Expert Systems with Applications. 2009;**36**(9):11864-11874

[107] Cha M, Haddadi H, Benevenuto F, Gummad KP. Measuring user influence on twitter: The million follower fallacy. In: 4th Int'l AAAI Conference on Weblogs and Social Media. Washington, DC; 2010

[108] Ebner M, Mühlburger H, Schaffert S, Schiefner M, Reinhardt W, Wheeler S. Getting granular on Twitter: Tweets from a conference and their limited usefulness for non-participants. In: Key Competencies in the Knowledge Society. Vol. 324. Boston: Springer; 2010. pp. 102-113

[109] Cuvelier E, Aufaure M-A. A buzz and e-reputation monitoring tool for twitter based on galois lattices. In: Andrews S, Polovina S, Hill R, Akhgar B, editors. Conceptual Structures for Discovering Knowledge, LNCS. Vol. 6828. Berlin: Springer; 2011. pp. 91-103

[110] Bonin F, Dell'Orletta F, Venturi G, Montemagni S. Contrastive filtering of domain-specific multi-word terms from different types of corpora. In: Proceedings of the workshop Multiword Expressions: from Theory to Applications (MWE 2010), 23rd International Conference on Computational Linguistics (COLING2010), Beijing, China, August 28. 2010. pp. 76-79

[111] Dell'Orletta F, Marchi S, Montemagni S, Plank B, Venturi G. The SPLeT-2012 shared task on dependency parsing of legal texts. In: Proceedings of the 4th Workshop on "Semantic Processing of Legal Texts" at LREC 2012. Istanbul, Turkey; 2012

[112] Petrov S, McDonald R. Overview of the 2012 shared task on parsing the web. In: Shared Task on Domain Adaptation for Parsing the Web At the First Workshop on Syntactic Analysis of Non-Canonical Language. At HLT-NAACL 2012 in Montreal on June 8, 2012. 2012

[113] Le Roux J, Foster J, Wagner J, Zadeh Kaljahi RS, Bryl A. DCUParis13 systems for the SANCL 2012 shared task. In: Notes of the First Workshop on Syntactic Analysis of Non-Canonical Language (SANCL). 2012

[114] Attardi G, Sartiano D, Simi M. Active learning for domain adaptation of dependency parsing on legal texts.

In: Proceedings of the 4th Workshop on “Semantic Processing of Legal Texts” at LREC 2012. Istanbul, Turkey; 2012

[115] Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on World wide web (WWW ‘10). New York, NY, USA: ACM; 2010. pp. 851-860

[116] Lanagan J, Smeaton AF. Using Twitter to detect and tag important events in sports media. In: Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona. Catalonia, Spain; 2011

[117] Leetaru KH. Culturomics 2.0: Forecasting large-scale human behavior using global news media tone in time and space. First Monday. 2011;**16**(9)

[118] Demir E, Fuhry D, Sriram B, Demirbas M, Ferhatosmanoglu H. Short text classification in twitter to improve information filtering. In: Proceedings of the ACM SIGIR 2010 Posters and Demos. Vol. 2010. Geneva, Switzerland

[119] Kucuktunc O, Cambazoglu BB, Weber I, Ferhatosmanoglu H. A large-scale sentiment analysis for Yahoo! answers. In: Proceedings of the fifth ACM international conference on Web search and data mining (WSDM ‘12). New York, NY, USA: ACM; 2012. pp. 633-642

[120] Zhai CX. Statistical Language Models for Information Retrieval (Synthesis Lectures Series on Human Language Technologies). Morgan & Claypool Publishers; 2008

[121] Archak N, Ghose A, Ipeirotis PG. Deriving the pricing power of product features by mining consumer reviews. Management Science. 2011;**57**(8):1485-1509

[122] Liu B, Hu M, Cheng J. Opinion observer: analyzing and comparing opinions on the Web. In: Proceedings of the 14th international conference on World Wide Web. 10-14 May 2005. Chiba, Japan; 2005

[123] Yuan J, Wang H, Xiao L, Zheng W, Li J, Lin F, et al. A formal study of shot boundary detection. IEEE Transactions on Circuits and Systems for Video Technology. 2007;**17**:168-186

[124] Xu C, Wang J, Lu H, Zhang Y. A novel framework for semantic annotation and personalized retrieval of sports video. IEEE Transactions on Multimedia. 2008;**10**:421-436

[125] Hauptmann AG, Yan R, Lin W-H, Christel MG, Wactlar H. Can high-level concepts fill the semantic gap in video retrieval? A case study with broadcast news. IEEE Transactions on Multimedia. 2007;**9**:958-966

[126] Stringa E, Regazzoni CS. Real-time video-shot detection for scene surveillance applications. IEEE Trans. on Image Processing. 2000;**9**(1):69-79

[127] Snoek CGM, Worring M. Concept-based video retrieval. Foundations and Trends in Information Retrieval. 2009;**4**(2):215-322

[128] Fan J, Elmagarmid AK, Zhu X, Aref WG, Wu L. ClassView: Hierarchical video shot classification, indexing and accessing. IEEE Transactions on Multimedia. 2004;**6**:70-86

[129] Tian Y, Hampapur A, Brown L, Feris R, Lu M, Senior A. Event detection, query, and retrieval for video surveillance. In: Ma Z, editor. Artificial Intelligence for Maximizing Content Based Image Retrieval. 2009. pp. 342-370

[130] Doretto G, Sebastian T, Tu P, Rittscher J. Appearance-based person reidentification in camera networks: problem overview and

current approaches. Journal of Ambient Intelligence and Humanized Computing. 2011;**2**:127-151

[131] Heisele B, Ho P, Poggio T. Face recognition with support vector machines: Global versus component-based approach. In: Proc. of the Eighth IEEE International Conference on Computer Vision. Vancouver, Canada; Vol. 2. 2001. pp. 688-694

[132] Candès EJ, Wakin MB. An introduction to compressive sampling. IEEE Signal Processing Magazine. 2008;**25**(2):21-30

[133] Baker JM, Deng L, Glass J, Khudanpur S, Lee C-H, Morgan N, et al. Research developments and directions in speech recognition and understanding, Part 1. IEEE Signal Processing Magazine. 2009;**26**(3):75-80

[134] Walker MA, Rudnicky A, Aberdeen J, Bratt EO, Garofolo J, Hastie H, et al. DARPA communicator evaluation: Progress from 2000 to 2001. In: ICSLP 2002. Vol. 1. 2002. pp. 273-276

[135] Hermansky H, Morgan N. RASTA processing of speech. IEEE Transactions on Speech and Audio Processing. 1994;**2**(4):578-589

[136] Saon G, Chien J-T. Special issue on fundamental technologies in modern speech recognition. IEEE Signal Processing Magazine. 2012:18-33

[137] Erzin E, Cetin AE. Interframe differential coding of line spectrum frequencies. IEEE Transactions on Speech and Audio Processing. 1994;**2**(2):350-352

[138] NIST: TRECVID video retrieval evaluation—Online proceedings 2002-2018. Available from: http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html

[139] Satta R, Fumera G, Roli F. Fast person re-identification based on dissimilarity representations. Pattern Recognition Letters. 2012;**33**(14):1838-1848

[140] Dee HM, Cohn AG, Hogg DC. Building semantic scene models from unconstrained video. Computer Vision and Image Understanding. 2012;**116**(3):446-456

[141] Abate A, Riccio MND, Tortora G. An ifs based approach for face recognition. In: Proc. IEEE International Conference on Image Processing. Vol. II. 2005. pp. 938-941

[142] Arandjelovi O, Cipolla R. An information-theoretic approach to face recognition from face motion manifolds. Image Vision Comput. 2006;**24**(6):639-647

[143] Beymer D, Poggio T. Face recognition from one example view. Tech. Rep. 1536. MIT AI Lab.; 1995

[144] Distasi R, Nappi M, Tucci M. Fire: Fractal indexing with robust extensions for image databases. IEEE Transactions on Image Processing. 2003;**12**(3):373-384

[145] Gao Y, Leung M. Face recognition using line edge map. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;**24**(6):764-779

[146] Gao Y, Leung M, Hui S, Tananda M. Facial expression recognition from line-based caricatures. IEEE Transactions on Systems, Man, and Cybernetics Part A. 2003;**33**(3):407-412

[147] Perner P. Prototype-based classification. Applied Intelligence. 2008;**28**(3):238-246

[148] Perner P, Attig A. Prototype-based classification for automatic knowledge acquisition of pathological processes at the cellular level. Transactions on Mass-Data Analysis of Images and Signals. 2010;**2**(1):41-54

[149] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces. In: Computer Graphics Proceedings SIGGRAPH'99. 1999. pp. 187-194

[150] Chowdhury AKR, Chellappa R. Face reconstruction from monocular video using uncertainty analysis and a generic model. Computer Vision and Image Understanding. 2003;**91**(1-2):188-213

[151] Cristinacce D, Cootes TF. Feature detection and tracking with constrained local models. In: Proceedings IEEE British Machine Vision Conference. 2006

[152] Georghiades AS, Belhumeur PN, Kriegman DJ. From few to many: Illumination cone models for face recognition under variable lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001;**23**(6):643-660

[153] Tan T, Yan H. Face recognition using the weighted fractal neighbor distance. IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews. 2005;**35**(4):576-582

[154] Brunelli R, Poggio T. Face recognition through geometrical features. In: LNCS. Vol. 588. Springer; 1992. pp. 792-800

[155] Fishler M, Elschlager R. The representation and matching of pictorial structures. IEEE Transactions on Computers. 1973;**C-22**(1):67-92

[156] Brunelli R, Poggio T. Face recognition: Features versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993;**15**(10):1042-1052

[157] Nefian AV, Hayes MH. Hidden Markov models for face recognition. In: Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Seattle, Washington, USA; 1998. pp. 2721-2724

[158] Pentland A, Moghadam B, Starner T. View-based and modular eigenspaces for face recognition. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1994. pp. 84-91

[159] Bartlett MS, Movellan JR, Sejnowski TJ. Face recognition by independent component analysis. IEEE Transactions on Neural Networks. 2002;**13**(6):1450-1464

[160] Wiskott L, Fellous J, Kruger N, von der Malsburg C. Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997;**19**:775-779

[161] Srisuk S, Petrou M, Kurutach W, Kadyrov A. Face authentication using the trace transform. In: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, Wisconsin, USA; 2003. pp. 305-312

[162] Saunders J. One of the most indicative and robust measures to discern voiced speech is the average zero-crossing rate (ZCR) of the time domain waveform. In: Real-time Discrimination of Broadcast Speech/Music. IEEE International Conf. On Acoustics, Speech, and Signal Processing (ICASSP). 1996

[163] Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE. 1989;**77**(2):257-286

[164] Perner P. Data reduction methods for technological industrial robots with direct teach-in-programming. Dissertation IH Mittweida 1985. 2nd edn. Fockendorf: IBAI Publishing; 2010. ISBN 978-3-940501-16-5

[165] Perner P, Attig A, Machno O. Novel method for the interpretation of spectrometer signals based on delta-modulation and similarity determination. Transactions on Mass-Data Analysis of Images and Signals. 2011;**3**(1):3-14 and The Patent: P. Perner "Method and Device for Automatically Determining a Substance Based on Spectroscopic Examinations," US020110153227A1

[166] Andén J, Mallat S. Deep scattering spectrum. IEEE Transactions on Signal Processing. 2014;**62**(16):4114-4128

[167] Jabloun F, Cetin AE, Erzin E. Teager energy based feature parameters for speech recognition in car noise. IEEE Signal Processing Letters. 1999;**6**(10):259-261

Chapter 6

Data Collection Techniques for Forensic Investigation in Cloud

Thankaraja Raja Sree and Somasundaram Mary Saira Bhanu

Abstract

Internet plays a vital role in providing various services to people all over the world. Its usage has been increasing tremendously over the years. In order to provide services efficiently at a low cost, cloud computing has emerged as one of the prominent technologies. It provides on-demand services to the users by allocating virtual instances and software services, thereby reducing customer's operating cost. The availability of massive computation power and storage facilities at very low cost motivates a malicious individual or an attacker to launch attacks from machines either from inside or outside the cloud. This causes high resource consumption and also results in prolonged unavailability of cloud services. This chapter surveys the systematic analysis of the forensic process, challenges in cloud forensics, and in particular the data collection techniques in the cloud environment. Data collection techniques play a major role to identify the source of attacks by acquiring evidence from various sources such as cloud storage (Google Drive, Dropbox, and Microsoft SkyDrive), cloud log analysis, Web browser, and through physical evidence acquisition process.

Keywords: distributed denial of service attacks, digital forensics, network forensics, web forensics, cloud forensics, mobile forensics

1. Introduction

In today's world, users are highly dependent on the cyberspace to perform all day-to-day activities. With the widespread use of Internet technology, cloud computing plays a vital role by providing services to the users. Cloud computing services enable vendors (Amazon EC2, Google, etc.) to provide on-demand services (e.g., CPU, memory, network bandwidth, storage, applications, etc.) to the users by renting out physical machines at an hourly basis or by dynamically allocating virtual machine (VM) instances and software services [1–3]. Cloud computing moves application software and databases to large data centers, where the outsourcing of sensitive data and services is not trustworthy. This poses various security threats and attacks in the cloud. For instance, the attackers use employee login information to access the account remotely with the usage of cloud [4]. Besides attacking cloud infrastructure, adversaries can also use the cloud to launch an attack on other systems. For example, an adversary can rent hundreds of virtual machine (VM) instances to launch a distributed denial-of-service (DDoS) attack. A criminal can also keep secret files such as child pornography, terrorist documents, etc. in cloud storage to remain clean. To investigate such crimes involved in the cloud, investigators have to carry out forensic investigations in the cloud environment. This arises the need for cloud forensics, which is a subset of network forensics. Cloud forensics is an application of scientific principles, practices, and methods to reorganize the events through identification, collection, preservation, examination, and reporting of digital evidence [5]. Evidence can reside anywhere in the cloud and it is more complex to identify the traces located in the cloud server.

The advancement of new technologies, frameworks, and tools enables the investigator to identify the evidence from trusted third parties, that is, cloud service provider (CSP). There are numerous techniques in cloud forensics that arises on the basis of cloud service models and deployment models. In the Software as a Service (SaaS) and Platform as a Service (PaaS) models, the customer does not have any control of the hardware and they need to depend on CSP for collecting the evidence, whereas, in the case of Infrastructure as a Service (IaaS) model, customers can acquire the virtual machine (VM) image and logs.

The forensic examiner isolates the attacked system in the virtualized environment by segregating and protecting the information from a hard disk, RAM images, log files, etc. This evidence is analyzed based on the artifacts of the attack traces left by the attacker [6, 7]. The forensic investigator relies on finding a series of information such as where, why, when, by whom, what, and how attack has happened. This chapter details the challenges in cloud forensics and also details the data collection techniques in the cloud.

2. Types of forensics

The forensic process is initiated after the crime occurs as a post-incident activity. It follows a set of predefined steps to identify the source of evidence. It is categorized into five groups, namely digital forensics, network forensics, Web forensics, cloud forensics, and mobile forensics.

- **Digital forensics**: According to National Institute of Standards and Technology (NIST) standards, it is the application of science to the identification, collection, examination, and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody for the data.

- **Network forensics**: It identifies and analyzes the evidence from the network. It retrieves information on which network ports are used to access the information.

- **Web forensics**: It identifies the evidence from the user history, temporary log files, registry, chat logs, session log, cookies, etc. as digital crimes occur on the client side with the help of Web browser.

- **Cloud forensics**: It is the application of digital forensics in the cloud and it is a subset of network forensics. It is harder to identify evidence in cloud infrastructure since the data are located in different geographical areas. Some examples of evidence sources are system log, application log, user authentication log, database log, etc.

- **Mobile forensics:** It is the branch of digital forensics that identifies evidence from mobile devices. The evidence is collected from the mobile device as call history, SMS, or from the memory.

2.1 Cloud forensic process flow

The cloud forensic process flow is shown in **Figure 1**, which is described as follows:

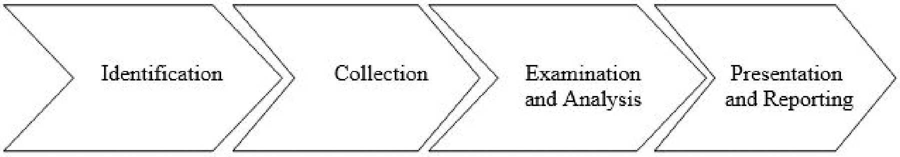


Figure 1.
Cloud forensic process flow.

- **Identification**: The investigator identifies whether crime has occurred or not.

- **Evidence collection**: The investigator identifies the evidence from the three different sources of cloud service model (SaaS, IaaS, and PaaS) [8]. The SaaS model monitors the VM information of each user by accessing the log files such as application log, access log, error log, authentication log, transaction log, data volume, etc. The IaaS monitors the system level logs, hypervisor logs, raw virtual machine files, unencrypted RAM snapshots, firewalls, network packets, storage logs, backups, etc. The PaaS model identifies the evidence from an application-specific log and accessed through API, patch, operating system exceptions, malware software warnings, etc.

- **Examination and analysis:** The analyst inspects the collected evidence and merges, correlates, and assimilates data to produce a reasoned conclusion. The analyst examines the evidence from physical as well as logical files where they reside.

- **Preservation:** The information is protected from tampering. The chain of custody has been maintained to preserve the log files since the information is located in a different geographical area.

- **Presentation and reporting:** An investigator makes an organized report to state his findings about the case.

3. Evidence collection

Evidence collection plays a vital role to identify and access the data from various sources in the cloud environment for forensic investigation. The evidence is no longer stored in a single physical host and their data are distributed across a different geographical area. So, if a crime occurs, it is very difficult to identify the evidence. The evidence is collected from various sources such as router, switches, server, hosts, VMs, browser artifacts, and through internal storage media such as hard disk, RAM images, physical memory, etc., which are under forensic investigation. Evidence is also collected through the analysis of log files, cloud storage data collection, Web browser artifacts, and physical memory analysis.

3.1 Cloud log analysis

Logging is considered as a security control which helps to identify the operational issues, incident violations, and fraudulent activities [9, 10]. Logging is mainly used to monitor the system and to investigate various kinds of malicious attacks. Cloud log analysis helps to identify the source of evidence generated from various devices such as the router, switches, server, and VM instances and from other internal components, namely hard disk, RAM images, physical memory, log files etc., at different time intervals. The information about different types of attacks is stored in various log files such as application logs, system logs, security logs, setup logs, network logs, Web server logs, audit logs, VM logs, etc., which are given as follows:

- *Application log* is created by the developers through inserting events in the program. Application logs assist system administrators to know about the situation of an application running on the server.

- *System log* contains the information regarding date and time of the log creation, type of messages such as debug, error, etc., system-generated messages related to the occurrence, and processes that are affected by the occurrence of an event.

- *Firewall log* provides information related to source routed packets, rejected IP addresses, outbound activities from internal servers, and unsuccessful logins.

- *Network log* contains detailed information related to different events that happened on the network. The events include recording malicious traffic, packet drops, bandwidth delays, etc. The network administrator monitors and troubleshoots daily activities by analyzing network logs for different intrusion attempts.

- *Web server log* records entries related to the Web pages running on the Web server. The entries contain history for a page request, client IP address, date and time, HTTP code, and bytes served for the request.

- *Audit log* records unauthorized access to the system or network in a sequential order. It assists security administrators to analyze malicious activities at the time of attack. The information in audit log files includes source and destination addresses, user login information, and timestamp.

- *VM log* records information specific to instances running on the VM, such as startup configuration, operations, and the time VM instance finishes its execution. It also records the number of instances running on VM, the execution time of each application, and application migration to assist CSP in finding malicious activities that happen during the attack.

Due to the increase in usage of network or new release of software in the cloud, there is an increase in the number of vulnerabilities or attacks in the cloud and these attacks are reflected in various log files. Application layer attacks are reflected in various logs, namely access log, network log, authentication log, etc., and also reflected in the various log file traces stored on Apache server. These logs are used for forensic examination to detect the application layer attacks. **Table 1** indicates the various attack information and the tools used for log analysis of different types of attacks. **Figure 2** shows the sample access log trace (**Table 2**).

- ***Sample Network Log Entry***

```
[**] [1:1407:9] SNMP trap udp [**] [Classification: Attempted Information
Leak] [Priority: 2] 03/12–15:14:09.082119 192.168.1.167:1052 - > 172.30.128.27:162
UDP TTL:118 TOS:0x0 ID:29101 IpLen:20 DgmLen:87.
```

- ***Sample Firewall Log Entry***

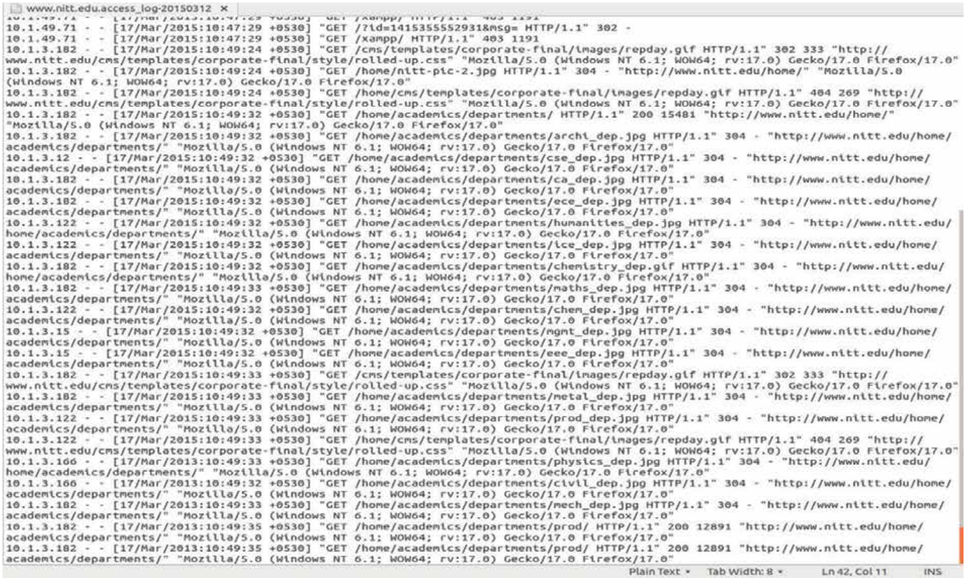
```
    03/12/2015 8:14:07 AM,"Rule ““Block Windows File Sharing”“ blocked
(192.168.1.54, netbios-ssn(139)).”,"Rule ““Block Windows File Sharing”“ blocked
(192.168.1.54, netbios-ssn(139)). Inbound TCP connection. Local address,service
is (KENT(172.30.128.27),netbios-ssn(139)). Remote address,service is
(192.168.1.54,39922). Process name is ““System”“.”
    03/12/2015 9:04:04 AM,Firewall configuration updated: 398 rules., Firewall
configuration updated: 398 rules.
```

- ***Sample Syslog Entries***

```
    Mar 1 06:25:43 server1 sshd[23170]: Accepted publickey for server2 from
172.30.128.115 port 21,011 ssh2.
    Mar 1 07:16:42 server1 sshd[9326]: Accepted password for murugiah from
10.20.30.108 port 1070 ssh2.
    Mar 1 07:16:53 server1 sshd[22938]: reverse mapping checking getaddrinfo for
ip10.165.nist.gov failed - POSSIBLE BREAKIN ATTEMPT!
    Mar 1 07:26:28 server1 sshd[22572]: Accepted public key for server2 from
172.30.128.115 port 30,606 ssh2.
    Mar 1 07:28:33 server1 su: BAD SU kkent to root on /dev/ttyp2.
    Mar 1 07:28:41 server1 su: kkent to root on /dev/ttyp2.
```

Types of log	Attacks	Tools for log analysis
DMesg log	This is not a log file, but this is used for determining anomalous activity from recent bots.	—
Debugging log	Stack tracing to determine the nature of application and service-based attacks.	—
Firewall log	Direct method for auditing the firewall.	Event Log Analyzer, event logging and monitoring services
System log	Determines if someone is trying or has executed buffer overflow.	Syslog-ng, Log & Event Manager
Network log	Determining Web-based attacks and DDoS attacks.	Splunk, Log4j2
Web server access log	Determining Web-based attacks (XSS, XSRF, SQLI), remote file inclusion, local file inclusion and flooding attacks.	Nihuo Web Log Analyzer
Web server error log	Determining Web-based attacks.	Nihuo Web Log Analyzer
VM log	Determining hypervisor-related attacks.	Virtual Machine Log Auditor, JVM controller
Authentication log	Auditing of attacks on credentials and determines the unauthorized access.	
Audit log	Determining unauthorized user access to the system and network. Includes destination addresses, user login information, and timestamp.	WP Security Audit Log, auditpol.exe
Database log	Determining database-related attacks.	Splunk, Nihuo Web Log Analyzer

Table 1.
Different types of logs, attacks, and the log analysis tool.



```
10.1.49.71 - - [17/Mar/2015:10:47:29 +0530] "GET /?id=1415355552931&msg= HTTP/1.1" 302 -
10.1.49.71 - - [17/Mar/2015:10:47:29 +0530] "GET /xampp/ HTTP/1.1" 403 1191
10.1.3.182 - - [17/Mar/2015:10:49:24 +0530] "GET /cms/templates/corporate-final/images/repday.gif HTTP/1.1" 302 333 "http://
www.nitt.edu/cms/templates/corporate-final/style/rolled-up.css" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:24 +0530] "GET /home/nitt-pic-2.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/" "Mozilla/5.0
(Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:24 +0530] "GET /home/cms/templates/corporate-final/images/repday.gif HTTP/1.1" 404 269 "http://
www.nitt.edu/cms/templates/corporate-final/style/rolled-up.css" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/ HTTP/1.1" 200 15481 "http://www.nitt.edu/home/"
"Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/archi_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.12 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/cse_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/ca_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/ece_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.122 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/humanities_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/
home/academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.122 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/ice_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/chemistry_dep.gif HTTP/1.1" 304 - "http://www.nitt.edu/
home/academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:33 +0530] "GET /home/academics/departments/maths_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.122 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/chem_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.15 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/mgmt_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.15 - - [17/Mar/2015:10:49:32 +0530] "GET /home/academics/departments/eee_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:33 +0530] "GET /cms/templates/corporate-final/images/repday.gif HTTP/1.1" 302 333 "http://
www.nitt.edu/cms/templates/corporate-final/style/rolled-up.css" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2015:10:49:33 +0530] "GET /home/academics/departments/metal_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.122 - - [17/Mar/2015:10:49:33 +0530] "GET /home/academics/departments/prod_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.122 - - [17/Mar/2015:10:49:33 +0530] "GET /home/cms/templates/corporate-final/images/repday.gif HTTP/1.1" 404 269 "http://
www.nitt.edu/cms/templates/corporate-final/style/rolled-up.css" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.166 - - [17/Mar/2013:10:49:33 +0530] "GET /home/academics/departments/physics_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/
home/academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.166 - - [17/Mar/2013:10:49:32 +0530] "GET /home/academics/departments/civil_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2013:10:49:33 +0530] "GET /home/academics/departments/mech_dep.jpg HTTP/1.1" 304 - "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2013:10:49:35 +0530] "GET /home/academics/departments/prod/ HTTP/1.1" 200 12891 "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
10.1.3.182 - - [17/Mar/2013:10:49:35 +0530] "GET /home/academics/departments/prod/ HTTP/1.1" 200 12891 "http://www.nitt.edu/home/
academics/departments/" "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:17.0) Gecko/17.0 Firefox/17.0"
```

Figure 2.
Sample access log trace as evidence.

S. No.	Fields	Value	Description
1	Remote Host	10.1.3.122	IP address of the HTTP user who makes HTTP resource request
2	Rfc931	—	Identifier used to determine client
3	Username	—	User name or user id used for authentication
4	Date: time Timezone	[17-Mar-2015: 10: 49: 33 + 530]	Date and timestamp of the HTTP request
5	HTTP request	GET/scripts/root.exe?/c+dir/HTTP/1.0	HTTP request containing (a) HTTP method—GET (b) HTTP request resource scripts/root.exe?/c+dir/ and (c) HTTP protocol version –1.0
6	Status code	200	Status of HTTP request, i.e., success or failure
7	Bytes	578	Number of bytes of data transferred during the HTTP request
8	Referral URL	https://www.nitt.edu/ OLCLD/view.php?q = book/	Referrer header of the HTTP request (containing URL of the page from which this request was initiated) if present, and "-" otherwise
9	User agent	Mozilla/4.08 [en] (Win98; I; Nav)	Browser Identification String

Table 2.
Description of the access log format.

3.2 Evidence collection from cloud storage

It is the process of collecting evidence from cloud storage such as Dropbox, Microsoft SkyDrive, Google drive, etc., using the Web browser and also by downloading files using existing software tools [11–13]. This helps to identify the illegal modification or access of cloud storage during the uploading or downloading of file contents in storage media and also checks whether the attacker alters the timestamp information in user's accounts. The Virtual Forensic Computing (VFC) tool is used by forensic investigators to identify evidence from VM image file. The evidence is accessed for each account using the Web browser running in the cloud environment by recording the encoded value of VM image. The packets are captured using network packet tools, namely Wireshark, snappy, etc., of each VM instance running in hosts. The account information is synchronized and downloaded using client accessing software of each device which is used to identify the source of evidence. The evidence is isolated from the files found in VM using "C:\Users\[username]\ Dropbox\" for Dropbox as shown in **Figure 3**. The zip file contains the name of the folder that can be accessed via the browser to determine the effect of a timestamp in a drive. If an attacker modifies the contents of a file, the evidence is found by analyzing the VM hard drive, history of files stored in the cloud, and also from a cache. It can also be analyzed by computing the hash value of the VM image. The evidence of Google Drive cloud storage is depicted in **Figure 4**.

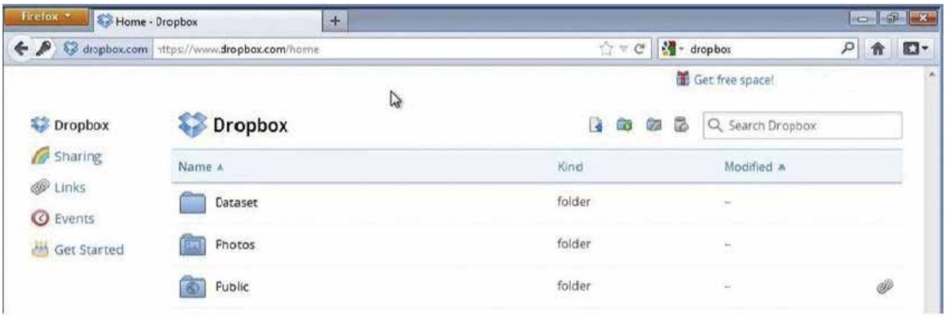
3.3 Evidence collection via a Web browser

The clients communicate with the server in the cloud environment with the help of a Web browser to do various tasks, namely checking email and news, online shopping, information retrieval, etc. [14–18]. Web browser history is a critical source of evidence. The evidence is found by analyzing the URLs in Web browser history, timeline analysis, user browsing behavior, and URL encoding, and is recovered from deleted information. Here is an example of Web browser URLs,

https://www.nitt.edu/en#files:/Documents/<Folder name>,

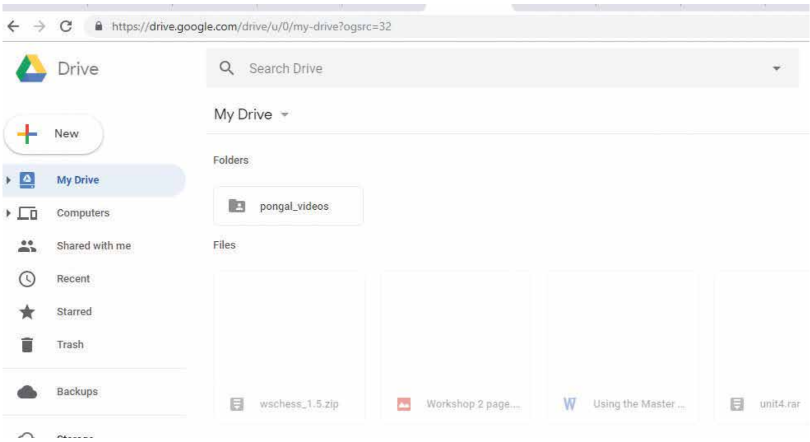
https://www.nitt.edu/en#files:/E:<Folder ID>.

Similarly, the evidence stored in Web browser cache at the root directory of a Web application is used to identify the source of an attack. **Table 3** indicates the evidence collection process and recovery method for various Web browsers.



Name ▲	Kind	Modified ▲
Dataset	folder	-
Photos	folder	-
Public	folder	-

Figure 3.
Dropbox evidence.



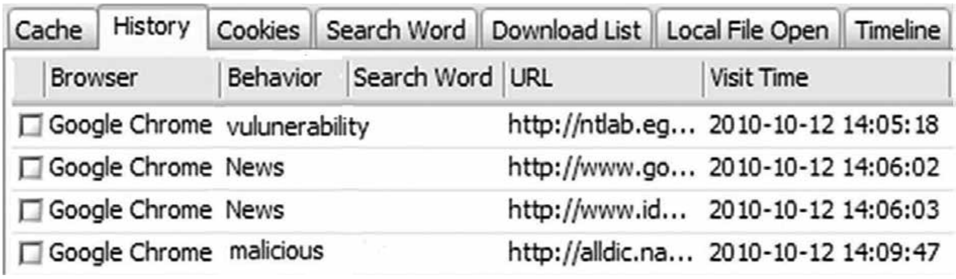
https://drive.google.com/drive/u/0/my-drive?ogsrc=32

Figure 4.
Google Drive evidence.

Web browser	Information to be analyzed	Tools for forensic investigation	Recovery method for evidence identification
Internet Explorer	Index.dat History Cache Cookies	Pasco Web historian 6.13 Index.dat analyzer 2.5 Net analysis 1.52 Encase 6.3 FTK 3.3 WEFA	Recovery from internet files Analyzing the index.dat files weekly/daily history Recovery of the evidence from index.dat file through carving method Recovery from cookies
Google Chrome	Bookmark history Bookmark downloads Cookies List of search words Cache	Chrome analysis 1.0 Net analysis 1.52 Cache back 3.17 WEFA	Recovery of session file through carving method
Mozilla Firefox	History Cookies history Download list Cache Bookmarks	Firefox forensic 2.3 Net analysis 5.2 Cache back 3.17 Encase 6.3 FTK 3.3 WEFA	Recovery of cache files
Safari	History Cache Cookies	Web historian 6.13 Net analysis 1.52 Cache back 3.17 Encase 6.3 FTK 3.3 WEFA	Recovery of session files, cookies
Opera	History Cache Cookies Bookmarks	Web historian 6.13 Net analysis 1.52 Cache back 3.17 Encase 6.3 WEFA	Recovery of cookies

Table 3.
Evidence collection process and recovery method for different Web browsers.

Here is an example of a Chrome forensic tool that captures and analyzes data stored in Google Web browser. It analyzes the data from the history, web logins, bookmarks, cookies, and archived history. It identifies the evidence from C:\Users\USERNAME\Appdata\Local\Google chrome\UserData\Default. **Figure 5** depicts the Google Chrome analysis forensic tool.



Browser	Behavior	Search Word	URL	Visit Time
□ Google Chrome	vulunerability		http://ntlab.eg...	2010-10-12 14:05:18
□ Google Chrome	News		http://www.go...	2010-10-12 14:06:02
□ Google Chrome	News		http://www.id...	2010-10-12 14:06:03
□ Google Chrome	malicious		http://alldic.na...	2010-10-12 14:09:47

Figure 5.
Chrome forensic analysis tool.

Forensic analysis framework	Evidence collection for cloud storage	Evidence collection for cloud log analysis
Evidence identification	Identification of evidence from cloud storage (Dropbox, iCloud, SkyDrive and Google Drive, etc.) and also from user account information	Identification of evidence from cloud log files
Evidence collection	Collecting the evidence from VM image to access the cloud storage account, using packet analysis tools such as Ethernet cap, Wireshark tool, Burp suite, etc. to capture packets between the client and server Collecting evidence from VM browser such as Google Chrome, chromium browser, Internet Explorer, Apple Safari, Mozilla Firefox, etc. Collecting the evidence from cloud storage namely, user account and password Collecting the evidence from client software to access the VM hard drive and also to synchronize the user account to retrieve the files and folders in VMs	Collecting the evidence from various sources in VM as log files, namely network log, access log, authentication log, error log, database log, etc. and through network analysis tools such as Wireshark, Snort, Snappy tool, Burp Suite, etc.
Evidence analysis	Identifying patterns from the evidence collection process to determine the source of attacks in cloud environment	Determining the attack patterns from cloud log files and analyzing these patterns using cloud traceback mechanism to identify the source of evidence.
Evidence presentation and reporting	Forensic investigator examines the evidence and presents the evidence in court	Identifying the evidence from analysis and reporting the evidence

Table 4.
Evidence collection process for cloud forensics.

3.4 Physical memory analysis

This has the ability to provide caches of cloud computing usage that can be lost without passive monitoring such as network socket information, encryption keys, and in-memory database. They are analyzed from the physical memory dump using the “pslist” function, which recovers the process name, process identifier, parent process identifiers, and process initiation time. The processes can be differentiated using the process names ©exe© on the Windows, and ©sync© on the Ubuntu and Mac OS. **Table 4** indicates the evidence collection process for cloud forensics in cloud storage and cloud log analysis.

4. Cloud forensics challenges

This section elucidates the forensic challenges in private and public cloud. It is observed from the literature that most of the challenges are applicable to the public cloud while fewer challenges are applicable to the private cloud environment.

4.1 Accessibility of logs

Logs are generated in different layers of the cloud infrastructures [2–7]. System administrators require relevant logs to troubleshoot the system, developers need logs for fixing up the errors, and forensic investigators need relevant logs to investigate the case. With the help of an access control mechanism, the logs can be acquired from all the parties, that is, from a user, CSP, and forensic investigator.

4.2 Physical inaccessibility

The data are located in different geographical areas of the hardware device. It is difficult to access these physical access resources since the data reside in different CSPs and it is impossible to collect the evidence from the configured device. If an incident occurs, all the devices are acquired immediately in case of a private cloud environment since an organization has full control over the resources. The same methods cannot be used to access the data in case of a public cloud environment.

4.3 Volatility of data

Data stored in a VM instance in a cloud will be lost when the VM is turned off. This leads to the loss of important evidence such as syslog, network logs, registry entries, and temporary Internet files. It is important to preserve the snapshot of the VM instance to retrieve the logs from the terminated VMs. The attacker launches an attack and turns off the VM instance, hence these traces are unavailable for forensic investigation.

4.4 Identification of evidence at client side

The evidence is identified not only in the provider’s side but also the client side. The user can communicate with the other client through the Web browser. An attacker sends malicious programs with the help of a Web browser that communicates with the third parties to access the services running in the cloud. This, in turn, leads to destroying all the evidence in the cloud. One way of collecting the evidence is from the cookies, user agent, etc., and it is difficult to obtain all the information since the client side VM instance is geographically located.

4.5 Dependence of CSP trust

The consumers blindly depend on CSPs to acquire the logs for investigation. The problem arises when CSPs are not providing the valid information to the consumer that resides in their premises. CSPs sign an agreement with other CSPs to use their services, which in turn leads to loss of confidential data.

4.6 Multitenancy

In cloud infrastructures, multiple VMs share the same physical infrastructure, that is, the logs are distributed across various VMs. The investigator needs to show the logs to court by proving the malicious activities occurring from the different service providers. Moreover, it also preserves the privacy of other tenants.

4.7 Decentralization

In cloud infrastructures, the log information is located on different servers since it is geographically located. Multiple users' log information may be collocated or spread across several layers and tiers in the cloud. The application log, network log, operating system log, and database log produce valuable information for a forensic investigation. The decentralized nature of the cloud brings the challenge for cloud synchronization.

4.8 Absence of standard format of logs

Logs are available in heterogeneous formats from different layers of a cloud at CSP. The logs provide information such as by whom, when, where, and why some incidents occurred. This is an important bottleneck to provide a generic solution for all CSPs and all types of logs. **Table 5** indicates the survey of literature that deals with the challenges of cloud forensics mainly for evidence collection process.

Authors	Discussion	Forensic process
Sang et al.	Log accessibility for SaaS & PaaS	Evidence collection
Zawood et al.	Focus on the integrity of log files	Evidence collection
Dystra et al.	Log collection and accessibility of logs	Evidence collection
Thorpe et al.	VM kernel logs for forensic investigation	Log contention
Boeck et al.	Confidentiality and log integrity	Evidence collection
Zaferulla et al.	Uses Eucalyptus logs for forensic investigation	Evidence analysis
Marty et al.	Collection of logs from different cloud components	Log retention
Sibiya et al.	Uses data mining techniques to collect logs for forensic investigation	Evidence collection
Patrascu et al.	Collection of specific logs	Evidence collection
Nakahara et al.	Evidence identification from different types of logs	Evidence collection and log retention

Table 5.
Challenges of cloud forensics.

5. Forensic tools

There are many tools to identify, collect, and analyze the forensic data for investigation. Juel et al. developed the PORs tool for the identification of online archives for providing integrity and privacy of files [19]. Dykstra et al. proposed a forensic tool for acquiring the cloud-based data in management plane [6]. It ensures trust in cloud infrastructures. Moreover, Encase and Access data FTK toolkit are used for the identification of trusted data to acquire the evidence. Similarly, tools such as evidence finder and F-response are used to find the evidence related to social networks. Dystra et al. proposed FROST, an open source OpenStack cloud tool for the identification of evidence from virtual disks, API logs, firewall logs, etc. [20].

6. Open research problems in cloud forensics

Many researchers have proposed various solutions to mitigate the challenges of cloud forensics. Some of the researchers have proposed new approaches to test the attacks in real-time environment. CSPs have not adopted the proposed solutions yet. Customers or investigators rely on CSPs to collect the necessary logs since they do not have direct physical access. Customers or investigators depend on CSP to collect the various information from the registry, hard disk, memory, log files etc. Even though various forensic acquisition process is proposed still the dependence of CSP remain unsolved. The critical issue is the usage of bandwidth resources. If the cloud storage is too high, then it results in more utilization of bandwidth. There is insufficient work evolved to preserve the chain of custody to secure provenance. There is no ideal solution for cybercrime scene reconstruction and preservation of evidence. Another critical issue is based on the modification of existing forensic tools that may lose evidence. Some researchers have proposed logging as a service to provide confidentiality, integrity, and authentication [3]. This solution is not suitable for IaaS cloud.

7. Case study

This section introduces a hypothetical forensic case study related to a cloud storage service and also describes a forensic investigation of the case.

7.1 Case study: cloud storage

The organization "X" found that their document named as "X_new.pdf" about the new release of a product has been leaked to their competitor [21–24]. "Mr. Morgan" was managing the credential files of the document stored in the cloud. At the initial stage of the investigation process, the suspect of the leaked file case was "Mr. Morgan." The forensic investigator has to identify the suspect by checking the organization network, or by the analysis of log files, or by collecting the trace of relevant file in the network. Mr. Morgan's network does not have any clue about the secrets since he uses only the personal computer (PC) and Android phone for business. To identify the suspects, the forensic investigator seized the PC and Android phone since these are the target devices used by the adversary. From the suspected devices, the leaked file has not been detected. Later, the investigator started analyzing the unallocated area in the file system, operating system, external devices such as hard drive, tablets, etc., and the Web service, but no evidence was found in the investigation. The investigator found that the Dropbox was installed in the PC and five files of config.db have been accessed recently. The forensic investigator issued the search warrant and identified the evidence in Dropbox by accessing Morgan's Dropbox storage with the username and password. It was observed that Morgan recently uploaded five files in Dropbox and identified that one of the files named as "XYZ_new.pdf" had the same contents as "X_new.pdf." Later, he deleted the traces of uploading or downloading the contents in PCs. The investigator found that Mr. Morgan has deleted the traces of the file contents and shared the evidence stored as "XYZ_new.pdf" with the competitor through an external SD card.

7.2 Case study: online railway ticket fraud

An online railway ticket booking service provider claimed that some unknown user had used the internet ticket booking facility to book 44 railway tickets using the stolen credit card details [25]. It has been charged back from the credit card companies for all transactions which led to a huge loss to the service provider. It is inferred from the investigation that the suspects have booked 44 tickets with different names of a person through the website at different locations. Later on, through investigation, the investigator found that the suspects arrived from a particular IP address, thereby seized the contents of the user accounts with the password, and the stolen credit cards were recovered from the suspects.

7.3 Case study: morphed photographs

The user got threatening pornographic emails from the adversary that one photograph was posted on the popular website [25]. The IP address for posting such threatening emails on the website was retrieved and was traced to a company. During an investigation, it was observed that the emails were sent from the company premises from one of the terminals. The log records and cookies were examined from the seized system and the morphed photographs were found in one of the systems used by the suspect. The mirror image of the hard disk was collected and analyzed using disk imaging and forensic analysis tools to recover all the data files required for the case. At the end of the investigation, it was found that the suspect was an ex-colleague of the company.

7.4 Case study: malicious insiders

Mr. X is an intruder who intends to exploit victims by sending malicious Web page in the cloud [26]. He uses a vulnerability to exploit the cloud presence of Buzz Coffee, a legitimate company. He installs a rootkit that injects a malicious payload into Web pages displayed and hides the malicious activity from the operating system. It redirects victims to the website, which infects them with malware. The users complain to the legitimate company that they are being infected, so the company went to the investigator to investigate the case by finding all the traces of the malicious Web page to identify the malicious user.

7.5 Case study: ransomware attack

A securities and brokerage firm became a victim of a ransomware attack [26]. The hacker demanded a ransom of two Bitcoins for each system that was infected. During the investigation process, it was observed that several other critical systems were infected with the same ransomware. Emails with malicious attachments appeared to be originating from a foreign location and were identified as the source of infection. The organization decided to take a proactive approach toward security with the focus on real-time monitoring to thwart such attacks in the future.

8. Conclusion

Cloud computing offers on-demand services (CPU, memory, network bandwidth, storage, applications, etc.) to users by allocating virtual instances and software services. Security is a major concern in the cloud wherein investigation of security attacks and crimes are very difficult. Due to the distributed nature of attacks and crimes in cloud, there is a need for efficient security mechanism. As cloud logs are spread across different virtual/physical machines (VM instances), switches, routers, etc., and also the customer (end user) is not aware of the activities of VM instances, cybercriminals exploit these sources to exhaust all the resources running in the cloud. Hence, evidence collection plays a crucial role to identify the suspects. However, collecting logs from the cloud infrastructure is extremely difficult because the investigator/security analyst has to depend on CSPs for collecting the logs and they have little control over the infrastructure. So, in order to identify the suspicious activity involved in the cloud, this chapter surveys the various forensic processes, evidence collection techniques for cloud forensics and the various challenges faced in cloud environment for forensic investigation.

Conflict of interest

The author does not have any conflict of interest.

Author details

Thankaraja Raja Sree* and Somasundaram Mary Saira Bhanu
Department of Computer Science and Engineering, National Institute of Technology, Tiruchirappalli, India

*Address all correspondence to: trajasree87@gmail.com

References

[1] Mell P, Grance T. The NIST Definition of Cloud Computing. Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, United States Department of Commerce. Gaithersburg. 2011:1-7 DOI: Special Publication 800-145

[2] Liu F, Tong J, Mao J, Bohn R, Messina J, Badger L, et al. NIST Cloud Computing Reference Architecture. Gaithersburg: NIST Special Publication. 2011. pp. 1-28. DOI: NIST SP 500-292

[3] Pichan A, Lazarescu M, Soh ST. Cloud forensics: Technical challenges, solutions and comparative analysis. Digital Investigation. 2015;**13**:38-57. DOI: 10.1016/j.diin.2015.03.002

[4] Guo H, Jin B, Shang T. Forensic investigations in cloud environments. In: 2012 International Conference on Computer Science and Information Processing (CSIP); 2012 Aug 24; IEEE. 2012. pp. 248-251. DOI: 978-1-4673-1411-4/12/

[5] Zawoad S, Hasan R. Cloud forensics: A meta-study of challenges, approaches, and open problems. 2013. arXiv preprint: 1302.6312

[6] Dykstra J, Sherman AT. Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques. Digital Investigation. 2012;**9**:S90-S98. DOI: 10.1016/j.diin.2012.05.001

[7] Marty R. Cloud application logging for forensics. In: Proceedings of the 2011 ACM Symposium on Applied Computing; 2011 Mar 21; ACM. 2011. pp. 178-184. DOI: 10.1145/1982185.1982226

[8] Anwar F, Anwar Z. Digital forensics for eucalyptus. In: 2011 Frontiers of Information Technology; 2011 Dec 19; IEEE. pp. 110-116. DOI: 10.1109/FIT.2011.28

[9] Khan S, Gani A, Wahab AW, Bagiwa MA, Shiraz M, Khan SU, et al. Cloud log forensics: Foundations, state of the art, and future directions. ACM Computing Surveys (CSUR). 2016;**49**(1):7. DOI: 10.1145/2906149

[10] Kent K, Souppaya M. Guide to Computer Security Log Management. Gaithersburg: NIST Special Publication; 2006. p. 92. DOI: N060928K

[11] Zhang OQ, Kirchberg M, Ko RK, Lee BS. How to track your data: The case for cloud computing provenance. In: 2011 Third IEEE International Conference on Cloud Computing Technology and Science; 2011 Nov 29; IEEE. -453. DOI: 446, 10.1109/CloudCom.2011.66

[12] Quick D, Choo KK. Forensic collection of cloud storage data: Does the act of collection result in changes to the data or its metadata? Digital Investigation. 2013;**10**(3):266-277. DOI: 10.1016/j.diin.2013.07.001

[13] Chung H, Park J, Lee S, Kang C. Digital forensic investigation of cloud storage services. Digital Investigation. 2012;**9**(2):81-95. DOI: 10.1016/j.diin.2012.05.015

[14] Nepal S, Ranjan R, Choo KK. Trustworthy processing of healthcare big data in hybrid clouds. IEEE Cloud Computing. 2015;**2**(2):78-84. DOI: 10.1109/MCC.2015.36

[15] Yusoff MN, Dehghantanha A, Mahmod R. Forensic investigation of social media and instant messaging services in Firefox OS: Facebook, Twitter, Google+, Telegram, OpenWapp, and Line as case studies. In: Contemporary Digital Forensic Investigations of Cloud and Mobile

Applications. 2017. pp. 41-62. DOI: 10.1016/B978-0-12-805303-4.00004-6

[16] Norouzizadeh Dezfouli F, Dehghantanha A, Eterovic-Soric B, Choo KK. Investigating social networking applications on smartphones detecting Facebook, Twitter, LinkedIn and Google+ artefacts on Android and iOS platforms. Australian Journal of Forensic Sciences. 2016;**48**(4):469-488

[17] Quick D, Choo KK. Google drive: forensic analysis of data remnants. Journal of Network and Computer Applications. 2014;**40**:179-193. DOI: 10.1016/j.jnca.2013.09.016

[18] Oh J, Lee S, Lee S. Advanced evidence collection and analysis of web browser activity. Digital Investigation. 2011;**8**:S62-S70. DOI: 10.1016/j.diin.2011.05.008

[19] Juels A, Kaliski BS Jr. PORs: Proofs of retrievability for large files. In: Proceedings of the 14th ACM conference on Computer and Communications Security; ACM. 2007. pp. 584-597. DOI: 10.1145/1315245.1315317

[20] Dykstra J, Sherman AT. Design and implementation of FROST: Digital forensic tools for the OpenStack cloud computing platform. Digital Investigation. 2013;**10**:S87-S95

[21] Chung H, Park J, Lee S, Kang C. Digital forensic investigation of cloud storage services. Digital investigation. Nov 1 2012;**9**(2):81-95. DOI: 1709/1709.10395

[22] Daryabar F, Dehghantanha A, Choo KK. Cloud storage forensics: MEGA as a case study. Australian Journal of Forensic Sciences. 2017;**49**(3):344-357. DOI: 10.1080/00450618.2016.1153714

[23] Martini B, Choo KK. Cloud storage forensics: ownCloud as a case study. Digital Investigation. 2013;**10**(4): 287-299. DOI: 10.1016/j.diin.2013.08.005

[24] Teing YY, Dehghantanha A, Choo KK, Yang LT. Forensic investigation of P2P cloud storage services and backbone for IoT networks: BitTorrent Sync as a case study. Computers & Electrical Engineering. 2017;**58**:350-363. DOI: 10.1016/j.compeleceng.2016.08.020

[25] http://prateek-paranjpe.blogspot.com/p/cyber-forensics-case-studies.html

[26] https://webforms.ey.com/Publication/vwLUAssets/ey-responding-to-cybercrimeincidents-in-india-new/$FILE/ey-responding-to-cybercrime-incidents-in-india.pdf

Chapter 7

Detectability of the Psychotropic Substance Cannabis in Head or Body Hair: Update of Forensic Criminalistics

Thorsten Floren M.A.

Abstract

Cannabis is a substance known and used by humans for thousands of years. Both from China and India, very old written traditions are known, which prove the use of cannabis in medicine. From the USA, there are opinions of scientists, who attribute a much higher risk for a lethal overdose to prescription drugs than is possible at all by the consumption of cannabis. Estimates for the area of the Federal Republic of Germany assume approximately 85,000 deaths per year due to the use of drugs and their undesirable side effects. The legal view of cannabis use is, with a few exceptions, a worldwide ban. In order to be able to implement these legal norms, investigators make use of a wide variety of procedures to prove the use of psychotropic substances in a court of law. In addition to the classical methods of taking blood samples and/or urine samples, it has also been possible for some time to prove the use of cannabis by examining hair samples. Only for about 40 years have various methods for retrospective forensic protection been available to criminalistics. Investigators use the special characteristics of hair, as opposed to blood, sperm or saliva. The hair grows more or less continuously from the root and stores substances from the metabolism of the body. This makes it possible for investigators to investigate the use of cannabis depending on the length and condition of the hair. In addition, the analytical methods for hair matrix analysis are continuously being developed. Studies have also shown various possibilities for the incorporation and accumulation of THC and its metabolites, the main active substances of cannabis, in hair. In addition to the classical substances for the preservation of evidence, further approaches result from scientific findings regarding the forensic analysis of hair. The securing and examination of head hair has become a standard procedure in investigations. It represents only a minor intervention for the test person. In addition, the storage of hair does not make any special demands on the temperature as an example. This makes hair a good examination tool for many parties involved in the procedure. However, scientific findings show the possibilities, risks and dangers of focusing exclusively on hair. In addition to the use of body hair as an object of investigation for the reliable collection of data, body hair can open up new possibilities. However, scientific findings must not be disregarded here. Every investigator must be aware at all times that he may only carry out interventions on accused persons on the basis of the respective state standards. The courts, too, can only give fair judgements if the investigating authorities have secured evidence that is also based on current scientific findings. In addition to the biological processes involved in the storage of foreign substances in the hair matrix, it is also necessary to describe in detail the various variants of analysis and evaluation on the various hair samples. All investigators must be aware at all times that they may only intervene against accused persons on the basis of the relevant state standards. The courts, too, can only pass just sentences if the investigating authorities have secured evidence based on the latest scientific findings. In addition to the biological processes involved in the storage of foreign substances in the hair matrix, it is also necessary to describe in detail the various variants of analysis and evaluation on the various hair samples.

Keywords: cannabinoid, THC, forensic head/body hair analysis, scientific findings

1. Introduction

The current relevance of an ever-increasing drug problem is illustrated here using the example of the Federal Republic of Germany. These data are of course not transferable one to one to every country. However, the numbers of trade and consumption of cannabis, for example, continue to rise throughout Europe [1].

According to police crime statistics, the number of trade offences involving cannabis in the Federal Republic of Germany increased by 18% between 2013 and 2017 [2]. The share of cannabis in comparison to other illegal drugs amounts to 64% and thus represents by far the largest share [3]. It should be mentioned here that crime statistics can only ever represent the so-called bright field. These are therefore only criminal acts that have become known to the police. This can be done by own investigations or by statements of witnesses. A frequently much larger proportion of criminal offences are in the dark field, are not known to the investigating authorities and are therefore not included in the statistical surveys.

The investigating authorities rely almost exclusively on the protection of head hair when securing evidence for the use of cannabis. This behaviour, which is partly voluntarily imposed or simply spread away by ignorance, disregards a large part of the securing material. In addition to the obvious hair on the head, body hair can also be secured on the arms, legs, armpits or skinned hair and used for evaluation.

However, it is also questionable to what extent hair from evidence is basically suitable for actual cannabis use?

2. Current status

2.1 Chemical structure of cannabis

Cannabis sativa L. is the Latin name for hemp [4]. In the chemical analysis of cannabis, more than 600 ingredients are already known. There are 100 cannabinoids and 50 hydrocarbons alone [5]. The dried leaves and flowers of the THC-rich cannabis varieties are called marijuana. The drug-typical THC content, which also contains psychoactivity, is between 1 and 20%. The leading cannabinoid responsible for this is the Δ9-THC [4]. The exact name of the Δ9-THC is Δ9-tetrahydrocannabinol. Due to the different numbering systems, the term Δ1-tetrahydrocannabinol is also used in part, but it names the same molecule. It is chemically $C_{21}H_{30}O_2$ and has a molecular weight of 314.47 Da [6] (**Figure 1**).

Recent research methods from the Δ9-THC have also analysed the main metabolites 11-Nor-9-carboxy-delta-9-terahydrocannabinol (CTHC)/(THC-COOH) and its glucuronide (CTHC-Glu). After cannabis use, CTHC-Glu is the main excretion product that can be detected in urine. However, studies on the detection of cannabis use in urine or blood have not provided comprehensive, reliable results of actual use [7]. The metabolite of the THC, 11-Nor-9-carboxy-delta-9-terahydrocannabinol (THC-COOH), can be stored in, among other things, by the supply in the hair bulb in the hair. THC-COOH is the most important main metabolite (**Figure 2**).

Another relevant stock is THCA-A (Δ9-tetrahydrocannabinolic acid A). This cannabinoid is the non-psychotoxic, biosynthetic precursor of THC and is present in fresh plant material of the cannabis plant [8] (**Figure 3**).

Furthermore, the cannabinoids cannabinol (CBN) and cannabidiol (CBD) are also analysed in hair analysis [9]. These substances are not psychoactive and are THC oxidation products.

2.2 Course of hair growth

The forensic preservation of evidence on hair is done by an individualising examination using molecular biological methods. In this way, incorporated foreign substances can be found and secured in the hair. In large parts of jurisprudence the view is taken that in principle a consumption, but also an existing abstinence of the consumption of cannabis can be determined. Here the possibility of retrospective analysis of cannabis use is advantageous. This can also be done for the past weeks up to months [10]. This is possible by a more or less continuous growth of the hair. The human hair grows about 1 cm per month. It allows an appropriate review (depending on the length of the hair) of a possibly long period of time by storing substances absorbed by the body in the hair [11]. This area of hair growth, known as the anagen

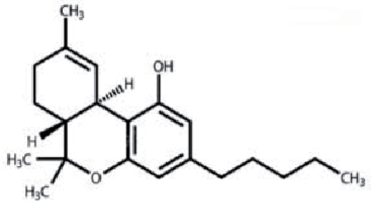


Figure 1.
Chemical formula of the THC.

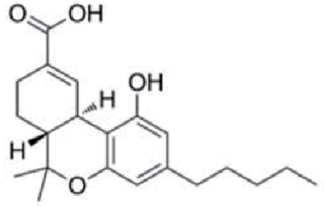


Figure 2.
Chemical formula of 11-Nor-9-carboxy-delta-9-terahydrocannabinol (THC-COOH).

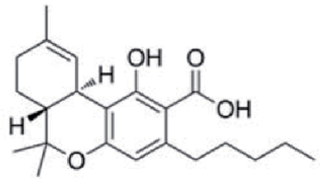


Figure 3.
Chemical formula of THCA-A.

phase, lasts about 4–6 years. About 80–95% of the hair in a healthy person is in this phase. In the second phase, the catagenic phase, cell division is stopped to form the hair. Hornification occurs. The second phase takes place over a period of 2 weeks. Only a few percent of the hair is in this stage of development at the same time. In the last, telogenic phase, the hair is already dead and is pushed out by a new hair forming in the hair root [12]. The diagram provides an overview of the structure and position of the hair root in the skin (**Figure 4**).

In the case of body hair, growth is significantly slower in a period between several months and a maximum of 1 year. The percentage of anagenic hair is only between 20 and 50% [14].

The storage of various substances, such as psychogenic substances in connection with hair growth and the associated pushing out of the hair, creates a temporal record of substance consumption in the hair in the manner of a tachograph, such as an example in a truck.

2.3 Storage of foreign substances

The hair absorbs foreign substances such as psychogenic substances in various ways. On the one hand, this occurs through contact with foreign substances in the body's bloodstream, which supplies the anagen hair with nutrients during growth. Through the metabolism and structure of the hair, the drug substances enter the resulting hair and are stored there [15].

The substances THC and THC-COOH can be stored in the hair in three ways. On the one hand a passive diffusion from blood capillaries directly into the hair matrix takes place. This takes place at the basement membrane of the hair follicle. The second way is the diffusion of sweat or sebum directly into the finished hair. The last possibility is contamination from the surrounding area [16]. The decisive factor is that THC-COOH is only formed in the body [17]. This also means that only THC-COOH can be actively incorporated into the hair matrix, which can be regarded as safe proof of cannabis use. However, the exact process of incorporating THC-COOH from the body into the hair matrix has not yet been fully researched scientifically [16].

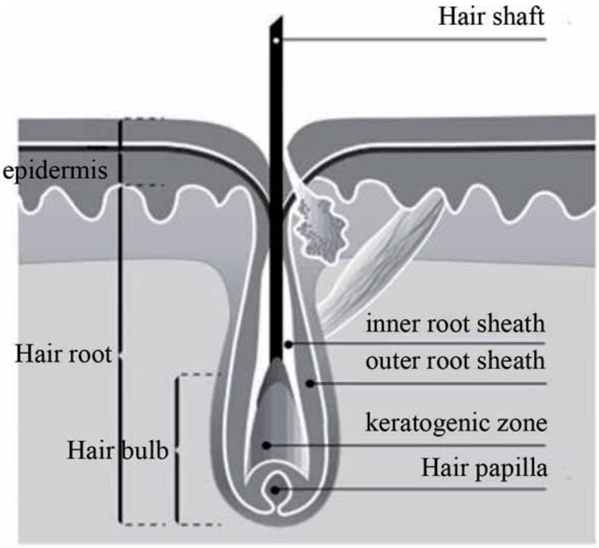


Figure 4.
Structure of the hair root in the skin [13].

In addition to the storage of substances, hair also absorbs them into the hair through external contact with substances and stores them there. The keratinized hair absorbs these foreign substances/substances, for example through drug-containing sweat, gases or dusts. The problem here is that in common laboratory practice, THC is still very often the only substance sought in the hair sample material. THC is found in a very high concentration in cannabis smoke and can lead to a strong contamination of the hair by an external build-up. The detection of THC-COOH in the laboratory is difficult and very time-consuming due to the very low dose [16].

In addition to this problem, very high concentrations of THCA-A (Δ9-tetrahydrocannabinolic acid A) also occur in chemical examinations of forensic hair samples [18]. THCA-A is the most important cannabinoid in fresh plant material. This substance is released, for example, by heating (smoking, baking). An uptake of THCA-A into the bloodstream could not be confirmed in earlier studies [8].

The concentration of the stored substances can, however, be drastically changed by environmental influences. The concentration is significantly reduced by the cosmetic treatment of the hair by bleaching, tinting or the form of a permanent wave [15]. The problem here is that there is no reliable data on how much the concentration decreases depending on external environmental influences. The degree of degradation can range from about 10% to almost 100% of the foreign substance [19]. Therefore, no scientifically reliable statement can currently be made as to which environmental influence causes which concentration change. These changes of the substance concentration can lead to a complete destruction of the hair substance and the foreign substances stored with it by a thermal treatment/thermal stretching of the hair [20]. Basically it can be stated that foreign substances are stored in the hair during the formation of new hair and thus in the anagen phase and can additionally be stored in the keratinized hair through external contact.

The difference, however, is that substances stored in the hair through growth fix themselves firmly in the hair matrix and then grow outwards from the hair root [14]. It takes about 10–14 days until the newly formed hair is outside the scalp [21].

3. Legal regulations

3.1 Police preservation of evidence of hair for the detection of cannabis

The example of the Federal Republic of Germany shows that the investigators there use hair on the head as evidence in addition to blood or urine in accordance with the current legal situation. These should be removed in the area "above the occipital protuberance." In accordance with the national requirements, corresponding documentation obligations are prescribed for the extraction in order to be able to bring the evidence into the criminal proceedings in a legally secure manner and ultimately present it as evidence in court.

Here again a central problem of the investigating authorities becomes apparent. Scientific findings are often only implemented or amended after a very long period of time. As a result, case law and police investigations lag behind scientific progress. This leads in large parts to a frequently delayed, fairer jurisdiction, since evidence is used in court which already no longer corresponds to the current scientific standard.

Closer cooperation between legislators, investigating authorities and researchers is urgently needed to improve the law.

4. Scientific status of the preservation of evidence on hair in cannabis use

Hair analysis is still a very young scientifically researched form of analysis. Forensic detection of drug use in hair has only been possible since 1979 [22]. The current state of science is a coupled analysis method. This starts with a separation of the substance mixture by liquid chromatography. This is also known as gas chromatography, which is used, for example, to determine the age of objects. This is followed by mass spectrometry. It is already possible to divide individual hairs into very short segments and examine them [23]. For this purpose, the hair must be washed before the examination in order to remove adhesions from the hair. It is precisely these adhesions that should not be included in the examination process, since only the substances stored in the hair are to be represented. The hair is then dried. They are then crushed and extracted. This extraction is necessary in order to break down the hair matrix with the stored substances. The final step is the analysis of the extracted material [24]. Here a high-quality measuring technique is needed. The sought-after substances are often only available in the nanogram or even in the picogram range [25].

The result obtained in this way is graphically processed and made available to the investigating authorities.

A major problem in the analysis and preparation of the sample material is the adhesion of substances to the hair. Especially the consumption of cannabis shows a significant problem due to its form of consumption. The active substance THC, which is present in cannabis, is most frequently consumed by smoking.

In addition, the state of scientific research in the field of hair analysis is very low. In 1995 a study was published which proved the storage of the THC metabolite THC-COOH via the blood circulation in the hair. However, this study was carried out on rats [26].

An exclusion, whether a contamination of the hair by sweat, saliva (coat care) etc. resulted, could not be produced. Despite this uncertainty, the detection of THC-COOH in the hair has been regarded as proof of cannabis use since that time [27].

In the forensic medicine department of the University Hospital Freiburg (Germany), a comprehensive study was conducted on the uptake and deposition of THC. In the first test set-up, subjects were exposed to passive smoke of cannabis for a period of 5 days in 3 weeks. The amount was a joint each time. A second experimental group took the active substance THC orally for 30 days (daily dose 7.5 mg). The third trial group only had to handle joints on 5 consecutive days. The study was designed in such a way that the hair of one experimental group only came into contact through cannabis smoke. The second group absorbed THC into their body and this opened up the possibility that THC could be incorporated into the hair matrix through hair growth. In this test group, body hair as well as head hair was examined. In the last test group, the hairs of the test subjects were apparently exposed to significantly less strain when handling cannabis joints, as they only emit a small amount of gas in a non-burning state.

In all three test groups hair samples were taken from the head after the end of the test series. The experimental group exposed to cannabis smoke showed a clearly detectable THC content of the hair. This value was particularly high in the area of the occipital protuberance. Thus exactly in the range, which is intended by the national police in the Federal Republic of Germany also for the removal. Other areas of the head hair showed significantly lower THC loads. The distribution of the THC deposition on the head was altogether very inhomogeneous. THCA-A was only detected in a very small amount in the hair.

In the second group of subjects who had taken the THC orally, no THC was detectable in the hair samples. This was not detectable both in the head and in body hair. However, THC-COOH could be detected in the hair samples. This was positively detectable both in the hairs of the test subjects. Also relatives of these test participants were analysed for THC-COOH. The hair of the relatives was also contaminated. It could be concluded that the oral intake of THC does not lead to storage of THC in the hair. However, THC-COOH was found in the hair. It was unusual that the area of the hair was contaminated with THC-COOH, which was proven to originate from a time before the study and therefore could not be caused by the orally ingested THC. The THC-COOH probably adhered to the hair via sweat and thus led to a positive result [28].

In the third group, which had only handled the joints, both THC exposure and THCA-A storage were observed. However, this was well below the levels of the group exposed to cannabis smoke. It is very problematic that even 1 month after the handling of the joints a positive result was still found in the analyses. Furthermore, the THCA-A value was higher than the THC value. In addition, the value was as high as for actual cannabis users. With the help of this study, the main route of THCA-A transmission into the hair could be shown, although actual use had not taken place [8].

In another study, the hair of children whose parents had been shown to use cannabis was analysed. Approximately 80 hair samples were examined, almost all of which were positive. The researchers found that the detected THCA-A and THC in the hair samples of children's hair was caused by contamination by hands or surfaces to which cannabis had previously been attached [29].

Another problem identified by a study is that the common analysis of hair by gas chromatography mass spectrometry followed by liquid-liquid extraction leads to THCA-A instability. This can lead to a partial conversion of THCA-A to THC, which leads to a wrong result in the evaluation [29].

Further comparative studies show a relatively high number of hairs tested negative, despite a positive urine test for cannabinoids [30].

The current methods of investigation used to provide evidence in court refer only to the analysis of THC. An examination according to THC-COOH, for example, is not carried out because the examination methods are very complex [28].

In addition to the problem of the correct choice of the examination material, there are further problems, especially in the area of hair analysis, such as the aging of the hair. External environmental influences such as solar radiation, cosmetic treatments, heat or heat influence further change the hair substance. The older the hair gets, the worse its evaluability becomes. Since hair is "youngest" at the root, it is generally not useful to analyse hair from a length of about 12 cm. The goal of an investigator with as long a hair as possible to be able to examine at the same time a long period of a test person's life is therefore not possible. At a distance of over 12 cm, the hair has been exposed to an average of so many external environmental influences that no more evaluable test results are to be expected [31].

In the comparison between head hair and body hair, the body hairs frequently show only a rather small length. This limits the evaluable period considerably more. But the lower influence of environmental influences clearly outweighs this disadvantage. Body hair is not treated so comprehensively with cosmetic products. But especially the area of thermal treatment (hairdryer, straightening iron, etc.) is almost completely omitted for body hair. The effect of heat has the greatest effect on the hair matrix and the foreign substances stored in it.

However, an area of the body hair should not be used for hair analysis. The area of the charm hair is often not to be used or used only very limitedly for analysis purposes. The contact with urine represents a very large impairment of the hair matrix. Due to the chemical composition and the properties of the urine, the evaluability of the hair decreases significantly. Also only very limited are axillary hairs suitable for a hair analysis. Contact with body sweat is not the biggest obstacle. The sometimes very frequent use of cosmetic products in the armpits puts so much strain on the hair that an analysis of the hair matrix often cannot provide meaningful results here either.

In principle, body hairs have a significantly shorter period of investigation or period of life due to their shorter length, but due to their position on the human body, which is partly protected against environmental influences, they represent a very valuable sample material.

It is currently problematic that there is only a very small number of available comparative studies with regard to the examination and analysis of body hair, especially in the area of growth rates and life cycle phases [32].

The presented studies clearly show how cautious a supposed positive THC, THC-COOH, or THCA-A content in a hair sample should be. If body hair had been examined in addition to head hair in this study, the result could have raised further questions. Many areas of the human body are covered with hair, even if most people are not always so aware of it. The only reason is that the hair is very thin, short and often so light that it does not stand out at first glance. A large part of this hair is covered by clothing most of the day and also at night. The area of swarm hair, armpit hair or hair on the stomach or back is exemplary. By covering these hairs with clothing, they are often strongly shielded from external environmental influences. A contamination for example by cannabis smoke or contact with cannabis products can either not take place at all or only very limitedly. The selection of the test material alone can have a considerable effect on the result. A sample of head hair exposed to cannabis smoke in the first experimental group showed a THC and clear THCA-A content.

A hair sample from the same person's back would probably have shown significantly lower or possibly no THC contamination.

The manifold possibilities for testing for THC, THCA-A or THC-COOH alone do not improve the situation. It is assumed that THC-COOH is incorporated into the hair matrix through the bloodstream. A positive test result would then also be a reliable result of actual cannabis use. Unfortunately, the scientific research situation is very low. This means that this storage route cannot be assumed with absolute certainty. In addition there is the difficulty of the complex analysis of hair according to THC-COOH. This is currently possible. However, due to the high costs and the costly examination method, it is not carried out area-wide. Rather, only THC is sought. However, the available studies clearly show that this research method does not stand up to any scientific evaluation. The oral intake of cannabis showed a negative THC result in the hair sample. When handling joints, the contamination with cannabis smoke and the examination of children's hair of cannabis users, positive results were obtained in the hair analysis, although there had actually been no consumption.

The investigation of THCA-A also did not give consistently correct results in hair analysis. The results are comparable to those of the THC study. False positive analyses were found even though no cannabis use was given.

5. Future perspectives

Current research clearly shows that it is precisely when cannabis is consumed, which normally occurs through inhalation, that by far the largest proportion of the THC content is deposited on the hair from the outside. There are now methods that can almost completely remove external adhesions from the hair. However, these methods are still very costly according to today's state of the art. The hair must be washed intensively and then the cleaned material must be examined for the detection of metabolically produced analytes. Due to the fact that the consumption of illegal drugs is on the rise and the proportion of cannabis is by far the highest, investigation procedures must be available for the investigating authorities, who can safely analyze mass crimes. The choice of sample material could lead to a significant increase in the quantitative and qualitative analysis of hair. The examination of body hair on head hair can significantly defuse the area of contamination with cannabis smoke. For this purpose, specific body hair must be taken, which is not exposed to direct contact with cannabis smoke due to its position, e.g., covered by clothing. In the field of hair analysis, however, it can be said that modern examination methods and the right selection of sample material can be used to make a statement on the consumption of, for example, cannabis drugs. However, the interpretation of the toxicological dose-concentration relationship determined is still difficult due to the different hair growth rate, external influences (age, colouring, heat, etc.), differences in metabolism and hair anatomy. They often do not allow a conclusive statement to be made about the amount of active ingredient actually absorbed. The low scientific knowledge regarding the substance to be investigated remains very problematic. Current hair samples are tested for THC on the basis of economic considerations. The studies clearly show that this way often does not lead to correct results. An examination method which would test for THC-COOH would presumably result in significantly better results.

6. Recommendations

Basically, however, it must be summarized that if a hair analysis is to be carried out according to scientific standards, extensive studies and test procedures are necessary beforehand. This is particularly important if you consider the subject of hair analysis from a legal point of view. The currently valid investigation according to the THC metabolite THC-COOH is from a scientific point of view not tenable as safe proof. Further studies must be carried out to prove the actual way of storing THC or its metabolites in the hair matrix. Here, the difference between the deposition by incorporation from the bloodstream of the THC and an attachment to the hair by smoke exposure, sweat, sebum, etc. must be shown. The difference between body hair and head hair must also be dealt with in the examinations. According to the current state of research, it is still not clear how THC is integrated into the hair matrix. In addition to the already investigated metabolites THC-COOH, the studies should also deal with other substances. In the end all approx. 600 ingredients of THC are known and a multitude of degradation products are added.

The problem of a false, positive THC result in a hair analysis has considerable consequences for the accused and also for the rule of law. It is therefore currently possible that a person who has only been exposed to passive cannabis smoke, which is not a criminal offence under national law, can face prosecution if only his hair is examined as evidence. Standing next to a person smoking a joint is not a criminal offence. However, if a positive result in a hair sample also suggests to the investigator that this person has used cannabis in the past although he did not do so at all, the innocent will be prosecuted.

7. Summary

Hair analysis has been available to investigating authorities since 1979 as a possible way of preserving evidence and evaluating evidence of cannabis use in criminal proceedings. The investigating authorities, using the example of the Federal Republic of Germany, have regulated the preservation of evidence for hair samples by means of ordinances. The occipital protuberance is explicitly named as the sampling point. This restriction of the sample location represents a clear restriction for an optimal preservation of evidence for the proof of cannabis use. The latest research results have shown that it is possible in a very complex laboratory procedure to extract and analyse only the THC-COOH content from the hair sample material, which was also incorporated into the hair matrix via the hair root and thus by the test person's body. Unfortunately, the research situation is still very low. There are still justified doubts whether the THC-COOH is incorporated into the hair matrix exclusively via the blood circulation or whether another way of contamination with THC or its metabolites is possible. The use of cannabis is only recorded in brightfield by police crime statistics. It is increasing successively and has become a mass crime. The possibility of an efficient investigation should therefore be used. THC accumulates on the hair, especially during inhaled consumption, due to the development of smoke. Studies have shown that, in addition to the adhesion by cannabis smoke, the inhalation into the hair can also occur through contact with sweat, fresh plant material and other surfaces. Both head hair and body hair are affected. The contamination pathways of THC, THC-COOH and THCA-A have been shown in various studies. This is illustrated again in the following diagram. Also the ways are represented, which are not yet finally clarified due to the current research situation (**Figure 5**).

In order to continue to use hair analysis as safe evidence in criminal proceedings, extensive studies should first be carried out to gain more insight into the actual route of enrichment and contamination of cannabis and its metabolites. The current state of scientific knowledge raises more questions than it can answer in relation to hair analysis. If, as an example, it continues to be confirmed that THC-COOH integrates into the hair matrix via the bloodstream after consumption, this would be a way of obtaining legally certain evidence.

Investigation authorities have a neutral legal mandate to secure, evaluate and analyse all incriminating and, in particular, exculpatory evidence. The current findings with regard to the analysis of head hair or body hair do not currently

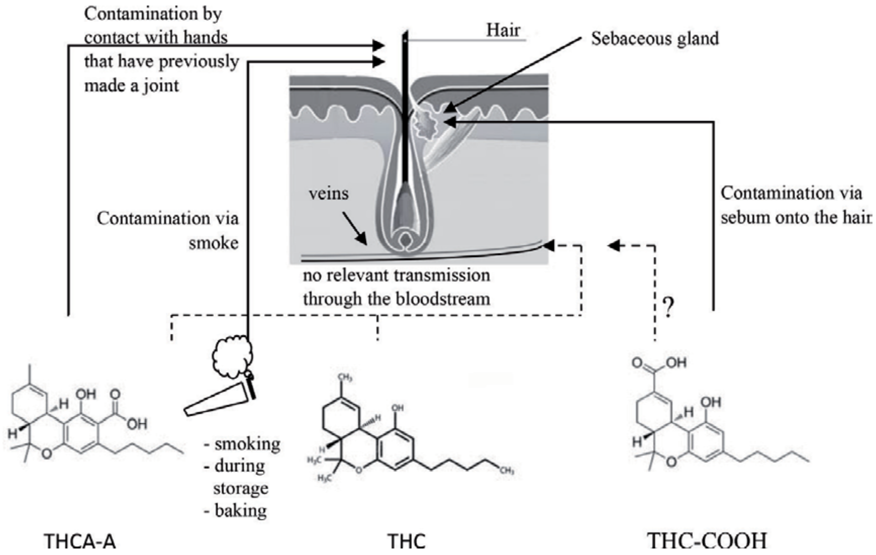

Contamination by contact with hands that have previously made a joint
Hair
Sebaceous gland
veins
Contamination via smoke
Contamination via sebum onto the hair
no relevant transmission through the bloodstream
?
- smoking
- during storage
- baking
THCA-A
THC
THC-COOH

Figure 5.
Incorporation of THC, THCA-A and THC-COOH into the hair.

represent a reliable, scientifically sound basis for the presentation of evidence. In addition to the hair sample analysis, the investigating authorities have the urine sample and the blood sample at their disposal as legally binding evidence. These should primarily be used for the analysis of cannabis use. This serves not only the accused, but also the rule of law and thus society.

Author details

Thorsten Floren M.A.
University of Police and Public Administration (HSPV NRW), Germany

*Address all correspondence to: thorsten.floren@web.de

References

[1] Musshoff F, Madea B. Review of biologic matrices (urine, blood, hair) as indicators of recent or ongoing cannabis use. Therapeutic Drug Monitoring. 2006;**28**(2):155

[2] Bundeslagebild. Rauschgiftkriminalität. BKA, Wiesbaden: Bundesrepublik Deutschland; 2017. p. 5. [Accessed: 23 May 2018]

[3] Bundeslagebild. Rauschgiftkriminalität. BKA, Wiesbaden: Bundesrepublik Deutschland; 2017. p. 6. [Accessed: 23 May 2018]

[4] Grotenhermen F. Cannabis und Cannabinoide, Pharmakologie, Toxikologie und therapeutisches Potenzial. 2. Auflage ed. Bern: Huber Verlag; 2005. p. 16

[5] Brauer P. Therapeutikum cannabis und post-polio-syndrom. Polio Europa aktuell. 2017;**72**:4

[6] Grotenhermen F. Cannabis und Cannabinoide, Pharmakologie, Toxikologie und therapeutisches Potenzial. 2. Auflage ed. Bern: Huber Verlag; 2005. p. 15

[7] Dietz L, Glaz-Sandberg A, Nguyen H, Mikus G, Aderjan R. Zur Kinetik des 11-Nor-9-carboxy-delta-9-tetrahydrocannabinol-glucuronids in Serum und Urin im Menschen. BA. 2007;**49**:171

[8] Moosmann B, Roth N, Auwärter V. Hair analysis for Δ9-tetrahydrocannabinolic acid A (THCA-A) and Δ9-tetrahydrocannabinol (THC) after handling cannabis plant material. Drug Testing and Analysis. 2016;**8**(1):128-132

[9] Paul R et al. Detection of cannabinoids in hair after cosmetic application of hemp oil. Scientific Reports. 2019;**9**(1):2582

[10] Madea B. Rechtsmedizin, Befunderhebung, Rekonstruktion, Begutachtung. Berlin: Springer Verlag; 2015. p. 644

[11] Wirth I. Kriminalistische Lexikon. 4. Auflage ed. Heidelberg: Kriminalistik Verlag; 2011. p. 265

[12] Scheufler F. Giften auf der Spur, Haare im Fokus der forensischen Analytik, Chemie in unserer Zeit. Weinheim: Wiley-VCH Verlag; 2018. p. 26

[13] WDR (Westdeutscher Rundfund): Quarks Scirpt, Script zur WDR-Sendereihe "Quarks & Co" Aachen: WDR Fernsehen; 1998. p. 7

[14] Fabian D, Baumgartner M, Koller M. In der Forensischen Toxikologie etabliert, in anderen Bereichen ist Vorsicht geboten, Sinn und Unsinn von Haaranalysen. Schweizerisches Medizin-Forum. 2016;**22**:467

[15] Madea B. Rechtsmedizin, Befunderhebung, Rekonstruktion, Begutachtung. Berlin: Springer Verlag; 2015. p. 645

[16] Moosmann B, Roth N, Auwärter V. Finding cannabinoids in hair does not prove cannabis consumption. Scientific Reports. 2015;**5**:1

[17] Grotenhermen F. Pharmacokinetics and pharmacodynamics of cannabinoids. Clinical Pharmacokinetics. 2003;**42**:327-360

[18] Auwärter V, Wohlfarth A, Traber J, Thieme D, Weinmann W. Hair analysis for Δ9-tetrahydrocannabinolic acid A—New insights into the mechanism of drug incorporation of cannabinoids into hair. Forensic Science International. 2010;**196**:10-13

[19] Baumgartner M. Nachweis des Konsums von psychotropen Substanzen und Alkohol mittels Haaranalyse. Therapeutische Umschau. 2011;**68**(5):269

[20] Fabian D, Baumgartner M, Koller M. In der Forensischen Toxikologie etabliert, in anderen Bereichen ist Vorsicht geboten, Sinn und Unsinn von Haaranalysen. Schweizerisches Medizin-Forum. 2016:468

[21] Fabian D, Baumgartner M, Koller M. In der Forensischen Toxikologie etabliert, in anderen Bereichen ist Vorsicht geboten, Sinn und Unsinn von Haaranalysen. Schweizerisches Medizin-Forum. 2016:469

[22] Baumgartner A, Jones P, Baumgartner W, Black C. Radioimmunoassay of hair for determining opiate-abuse histories. Journal of Nuclear Medicine. 1979;**20**(7):748

[23] Scheufler F. Giften auf der Spur, Haare im Fokus der forensischen Analytik, Chemie in unserer Zeit. Weinheim: Wiley-VCH Verlag; 2018. p. 28

[24] Baumgartner M. Nachweis des Konsums von psychotropen Substanzen und Alkohol mittels Haaranalyse. Therapeutische Umschau. 2011;**68**(5):270

[25] Baumgartner M. Nachweis des Konsums von psychotropen Substanzen und Alkohol mittels Haaranalyse. Therapeutische Umschau. 2011;**68**(5):272

[26] Nakahara Y, Takahashi K, Kikura R. Hair analysis for drugs of abuse. X. Effect of physicochemical properties of drugs on the incorporation rates into hair. Biological and Pharmaceutical Bulletin. 1995;**18**:1223-1227

[27] Kintz P. Analytical and Practical Aspects of Drug Testing in Hair. Boca Raton: Taylor & Francis Group; 2007

[28] Moosmann B. Neue Aspekte in der Haaranalytik auf Cannabinoide. Toxichem Krimtech. 2015;**82**(3):311

[29] Moosmann B, Roth N, Hastedt M, Jacobsen-Bauer A, Pragst F, Auwärter V. Cannabinoid findings in children hair—What do they really tell us? An assessment in the light of three different analytical methods with focus on interpretation of Δ9-tetrahydrocannabinolic acid A concentrations. Drug Testing and Analysis. 2015;**7**(5):349-357

[30] Musshoff F, Driever F, Lachenmeier K, et al. Results of hair analyses for drugs of abuse and comparison with self-reports and urine tests. Forensic Science International. 2006;**156**:311

[31] Scheufler F. Giften auf der Spur, Haare im Fokus der forensischen Analytik, Chemie in unserer Zeit. Weinheim: Wiley-VCH Verlag; 2018. p. 32

[32] Scheufler F. Giften auf der Spur, Haare im Fokus der forensischen Analytik, Chemie in unserer Zeit. Weinheim: Wiley-VCH Verlag; 2018. p. 31

Chapter 8

Novel Methods for Forensic Multimedia Data Analysis: Part II

Petra Perner

Abstract

We are proposing a new concept for new multimedia data processing techniques for varied multimedia sources. We address our work toward speech, video and images, handwriting and text documents. The methods and techniques will form a toolkit that can be used for different cases in a court of law in order to extract information from different multimedia sources. In addition to the data mentioned above, social media (e.g., Facebook and Twitter) provide multimedia data in new formats. Such data allow the investigator to identify and compare objects, events or persons based on data properties, including biometric features or more symbolic features that point to coincidences and anomalies. We continue our work of novel methods for forensic multimedia data analysis Part I by a description of related work and a proposal of the methods and techniques we are developing beyond the state of the art for handwriting, multimedia feature extraction, novelty detection, legal aspects and cloud computing. Then, we describe the tasks that should be solved by the system and the different multimedia data. We describe expected results of our proposed toolkit. Finally, we summarize the objective of our work.

Keywords: multimedia forensic data analysis, standardization of forensic data analysis, video and image enhancement, video analysis, image analysis, speech analysis, case-based reasoning, multimedia feature extraction, handwriting, twitter data analysis, novelty detection, legal aspects

1. Introduction

The objective of this work is to provide novel methods and techniques for the analysis of forensic multimedia data. These methods and techniques should form a novel toolkit for automatic forensic multimedia data. The data modalities the proposed work is considering are images and videos, text, handwriting, speech and audio signals, social media data, log data and genetic data. The integration of methods for all these different data modalities in one toolkit should allow the cross-analysis of these data and the detection of events by interlinking between these data. The proposed methods will face on standard forensic tasks, e.g., identification of events, persons or groups, and device recognition. Together with the end users and the police forces, new standard tasks will be worked out during the project and will give a new input to the standardization aspect of forensic data analysis.

The proposed novel methods and techniques will consider all aspects of multimedia data analysis such as device identification and trustworthiness of the data, signal enhancement, preprocessing, feature extraction, signal and data analysis and interpretation.

The chapter is continuation of Part I of novel methods for forensic multimedia data analysis [1]. The main aspects of multimedia forensic data analysis, the background and the system architecture are described over in Part I. Related work and the proposed methods and techniques that go beyond the state of the art for video, images, speech, multimedia feature extraction, text and Twitter data analysis are also described in Part I. Here we are continuing our description of related work and the proposed methods and techniques that go beyond the state of the art for handwriting, novelty detection, law aspects and cloud computing in Section 2. Task system has to be solved based on the different multimedia data, for case-based reasoning and legal aspects, which are explained in Section 3. In Section 4, we describe what kind of results we expect from our system and our work. Finally, we give conclusions in Section 5.

2. Related work to be continued from Part I

2.1 Handwriting recognition

2.1.1 State of the art

Criminologists and forensic document examiners have been investigating the clues to identify the handwritings for decades [2]. However, identification by human experts has the drawbacks of nonobjective measurements and nonreproducible decisions besides the cost of expert training. Recent attempts in computer-supported handwriting identification aim to address the challenge of doing this task using computer vision and pattern recognition techniques. The Forensic Information System for Handwriting (FISH) system developed by German law enforcement is followed by more recent ones including WANDA and CEDAR-FOX systems [3–5].

As an initial step to identification, the documents are required to be processed in order to handle the noise and for enhancement [6]. Then some similarity measures have to be defined to identify all the writings of a person as same and to differentiate it from others' writings [7]. The writer identification systems make use of the individuality of handwritings and try to capture the characteristics of handwriting in a macro level by the style of handwriting and in a micro level by the shapes of the characters.

The literature is mostly dominated by character-based systems. HCLUS prototype matching techniques [3], dynamic time warping-based techniques [8, 9] and structural features [10] have been recently proposed for matching allographs (prototypical character shapes). The disadvantage of character-based systems is the requirement for character segmentation and the character classification (whether the character is y or g, etc.) prior to finding the similarities and dissimilarities on a specific character. However, both tasks are challenging and error-prone when free handwriting is the issue. The shape of the characters may change with the size of writing or with pen size and type. The characters can be written differently in a different context, i.e., in different words, and that requires the consideration of the preceding and following characters. Most importantly, it is very difficult to apply those systems to large datasets.

2.1.2 Beyond the state of the art

When the documents in forensic applications are considered, the preprocessing step gains more importance due to the variety of artifacts and degradations that are generated intentionally to hide the evidences or happened because of the environmental conditions. The enhancement methods should be carefully applied in order not to remove important cues while reducing the noise. The methods that will be developed will be adaptive and will learn from a large number of data by incorporating the expert feedback when necessary.

Inspired by cognitive studies that have observed the human tendency to read whole words at a time [11], in recent studies word-spotting techniques have been proposed as an alternative to character-based systems in the word retrieval literature [12–16]. As a novel direction in forensic handwriting identification, to be able to work in the large scale, we will follow the direction in word spotting and describe and match the words rather than characters. Generic image features will be used to describe the word images as a whole. This will enable us to capture the writing habits and styles of individuals and also the character variations in a different context.

Going beyond the writer identification and verification, recognition of printed and handwritten text from documents (such as letters and notes) and the text in the photographs taken in the environment (such as the labels of shops, buildings or streets and advertisements on boards) will be another issue that will be considered. While the commercially available optical character recognition systems are very successful for printed documents, recognition of handwritten text continues to be a challenge [17, 18]. More importantly, the recognition of words in unconstrained settings or “in the wild” is still an interesting problem [19]. Besides scene and object detection and recognition methods that will be developed, recognition of text in the images, ranging from license plates to shop labels and even the text on clothes, will provide important cues for the identification of places.

Both identification and recognition will be attacked similar to generic object recognition, and word images will be described using advanced image descriptors to be able recognize and identify text in multi-author multi-language cases. Besides the SIFT [20]-based and k-AS [21]-based features that we considered in our previous studies [22–24], SURF [25], FREAK [26] and Shape Context [27]-based features will be adopted for word description. Together with the statistical analysis of the occurrence of some features, the spatial layout of extracted features will also be considered.

Large volumes of data will be exploited through data mining techniques in order to learn the in-class similarities and between-class differences. Novel similarity measures will be developed. The main goal will be to provide the sufficient amount of data to the experts to be validated. The main challenge will be not to miss any important data while reducing the huge volume. Therefore, the methods that will be developed for similarity should be both robust and fast. We will benefit from expertise of partners in different areas to design new similarity measures for handwriting matching.

2.2 Novelty detection

2.2.1 State of the art

The aim of novelty detection is to recognize inputs that cannot be properly represented by information provided by previous inputs (i.e., a nominal distribution). Recognizing that an input differs in some respect from previous inputs is a very important capability of a learning system. In classification problems novelty detection is particularly useful when a relevant class is under-represented in the data, so that a classifier cannot be trained to reliably recognize that class or when hierarchical classifiers trained on different concept information disagree on the output.

The goal of novelty detection is twofold: to be as accurate as possible in detecting inputs which do deviate from the nominal distribution (true positives) and to predict how many normal inputs will be erroneously flagged as positives (false positives). Novelty detection is also known as one-class classification [28] or learning from only positive (or only negative) examples. The standard approach has been to assume that novelties are outliers with respect to the nominal distribution and to build a novelty detector by estimating a level set of the nominal density. This approach allows fixing a threshold for acceptance of new data while having a degree of control over the number of false alarms raised. Using this framework, novelty detection can be interpreted as a binary classification problem. Several approaches have been utilized to tackle this problem: statistical methods, neural networks and support vector method approaches (see [29–32], for good reviews of these techniques). Bayesian methods have been used to provide a nonparametric estimation of the probability distribution [33], and content-based reasoning (CBR), based on Bayesian decision theory, has also been utilized. A common drawback of all these approaches is the assumption that novelties are uniformly distributed on the support of the nominal distribution, which is not true in most cases, mainly when the feature space dimension is high.

Novelty detection has been already applied with success on single modalities of the forensic multimedia data. For instance, the detection and classification of abnormal events in a surveillance video has been studied in [34–36]. In [33], novelty detection is applied in online document clustering, and in [37], novelty detection is applied on image sequences.

A new and promising approach to novelty detection in audiovisual data has been proposed in [38]. In this approach, the novelty detection is not the negative output of multiple classifiers but the disagreement of several concept hierarchical classifiers trained from different but hierarchically related concept. Here, the novelty is represented not by a fully new item but by relevant changes from previous seen items. In forensic multimedia data, this situation is very common when only one modality is affected.

According to [33], several factors make the novelty detection problem very challenging:

- The definition of a normal region that encompasses every possible normal behavior is very difficult. In addition, the boundary between normal and anomalous behavior is often not precise. Thus, an anomalous observation that lies close to the boundary can actually be normal and vice versa.

- When anomalies are the result of malicious actions, the malicious adversaries often adapt themselves to make the anomalous observations appear normal, thereby making the task of defining normal behavior more difficult.

- In many domains, normal behavior keeps evolving, and a current notion of normal behavior might not be sufficiently representative in the future.

- The exact notion of anomaly is different for different application domains. For example, the rate of change for novelty may be different in each application. Thus applying a technique developed in one domain to another one is not straightforward.

- The availability of labeled data for training/validation of models used by anomaly detection techniques is usually a major problem.

- Often, the data contain noise that tends to be similar to the actual anomalies and hence is difficult to distinguish and remove.

2.2.2 Beyond the state of the art

We will consider novelty detection as a CBR problem [34]. The CBR-based novelty detection will consist of successively adapting or evolving the previously obtained solutions, taking into account the data properties, the user's needs and any other prior knowledge into account. We will use a combination of statistical and similarity-based methods as the solution to the problems underlying the CBR methodology. Our proposed scheme differs from existing methodologies on novelty detection [29, 30, 39–41] since it can perform simultaneously novelty detection and handling and also considers the incremental nature of the data.

2.3 Legal aspects

2.3.1 State of the art

The normative compliance of the methodologies and tools proposed by the project will be assessed by reference to the European and national legal framework on data protection and privacy [42].

Data protection is a fundamental right in Europe, enshrined in Article 8 of the Charter of Fundamental Rights of the European Union as well as in Article 16(1) of the Treaty on the Functioning of the European Union (TFEU).

The central legislative instrument for the protection of personal data in Europe is the EU's 1995 Directive[1]. In order to come to a complete revision of the entire framework concerning data protection, more consistent with changes in the single market and with stronger needs to ensure security of European citizens, since 2009, the European Commission launched public consultations on data protection[2] and engaged in intensive dialog with stakeholders. On 4 November 2010, the Commission published a communication on a comprehensive approach on personal data protection in the European Union[3] that sets out the main themes of the reform. "After assessing the impacts of different policy options[4], the European Commission is now proposing a strong and consistent legislative framework across Union policies, enhancing individuals' rights, the Single Market dimension of data protection and cutting red tape for businesses"[5]. The Commission proposes that the new framework should consist of:

[1] Directive 95/46/EC on the protection of individuals with regard to the protection of personal data and on the free movement on such data, OJ L 281, 23.11.1995, p. 31.

[2] Two public consultations have been launched on the data protection reform: one from July to December 2009 (http://ec.europa.eu/justice/news/consulting_public/news_consulting_0003_en.htm) and a second one from November 2010 till January 2011 (http://ec.europa.eu/justice/news/consulting_public/news_consulting_0006_en.htm).

[3] COM(2010)609.

[4] The Impact Assessment SEC(2012)72.

[5] See p. 4, final COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Safeguarding Privacy in a Connected World A European Data Protection Framework for the 21st Century Brussels, 25.1.2012 COM(2012) 9.

- A Regulation[6] (replacing Directive 95/46/EC) setting out a general EU framework for data protection

- A Directive[7] (replacing Framework Decision 2008/977/JHA[8]) setting out rules on the protection of personal data processed for the purposes of prevention, detection, investigation or prosecution of criminal offenses and related judicial activities

By proposing a specific Directive regulating the use of personal data for criminal investigations, the EU legislator acknowledges the importance of innovating by means of a legislative instrument, the specific theme of personal data processing in criminal investigations, a theme that was left aside by the Directive 95/46 and only partially regulated by Framework Decision 2008 [43, 44]. It left to national legislations all the decisions about the legitimacy in processing personal data for purpose of crime detection, thus preventing a global European intervention against criminality.

The new Directive intends to make a distinction between the fundamental (but not "absolute") right to data protection and its social profile in the light of achieving a global security. An EU legislator states that "The right to the protection of personal data, enshrined in Article 8 of the Charter of Fundamental Rights and in Article 16(1) TFEU, requires the same level of data protection throughout the Union"[9] and that "According to the principle of subsidiarity (Article 5(3) TEU), action at Union level shall be taken only if and in so far as the objectives envisaged cannot be achieved sufficiently by Member States, but can rather, by reason of the scale or effects of the proposed action, be better achieved by the Union".

The need to respect the sovereignty principle explains why the EU chose the instrument of Directive that leaves a space of more flexibility to Member States, instead of the adoption of a Regulation, that would have a stronger impact on national legislation on privacy protection [45].

The purpose of the new rules is highlighted in Article 1: 'Subject matter and objectives. 1. This Directive lays down the rules relating to the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offenses or the execution of criminal penalties. 2. In accordance with this Directive, Member States shall: (a) protect the fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data; and (b) ensure that the exchange of personal data by competent authorities within the

[6] Brussels, 25.1.2012 COM(2012) 11 final 2012/0011 (COD) Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) {SEC(2012) 72 final} {SEC(2012) 73 final} SEC 2012 ...The Regulation also makes a limited number of technical adjustments to the e-Privacy Directive (Directive 2002/58/EC as last amended by Directive 2009/136/EC—OJ L 337, 18.12.2009, p. 11) to take account of the transformation of Directive 95/46/EC into a Regulation.

[7] COM(2012) 10: DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offenses or the execution of criminal penalties, and the free movement of such data.

[8] Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, OJ L 350, 30.12.2008, p. 60.

[9] COM 2012,10, Sections 3 and 3.2.

Union is neither restricted nor prohibited for reasons connected with the protection of individuals with regard to the processing of personal data.

2.3.2 Beyond the state of the art

The main innovations concern the following:

- The introduction of the definitions of 'personal data breach,' 'genetic data' and 'biometric data,' 'competent authorities' (based on Article 87 TFEU and Article 2(h) of Framework Decision 2008/977/JHA) and, of a 'child,' based on the UN Convention on the rights of the child

- The distinction between different categories of data subjects (art.5): 'Member States shall provide that, as far as possible, the controller makes a clear distinction between personal data of different categories of data subjects, such as:

    a. persons with regard to whom there are serious grounds for believing that they have committed or are about to commit a criminal offense;

    b. persons convicted of a criminal offense;

    c. victims of a criminal offense, or persons with regard to whom certain facts give reasons for believing that he or she could be the victim of a criminal offense;

    d. third-parties to the criminal offense, such as persons who might be called on to testify in investigations in connection with criminal offenses or subsequent criminal proceedings, or a person who can provide information on criminal offenses, or a contact or associate to one of the persons mentioned in (a) and (b); and

    e. persons who do not fall within any of the categories referred to above.'

2.4 Cloud computing

2.4.1 State of the art

Many of the tasks, such as video analysis, text mining or case-based reasoning, are very compute and data intensive. Processing large amounts of multimedia data for police investigation purposes must take into account scalability and performance requirements in order to be usable in a professional context. Cloud computing [46] has emerged as a good model for providing compute and storage resources on demand. To benefit from these available resources, applications that are deployed in the cloud need to be designed to scale and to be able to benefit from parallel processing of data with approaches such as map-reduce [47].

2.4.2 Beyond the state of the art

In this proposal, we will advance the state of the art by designing algorithms for video-analysis, text mining or case-based reasoning that are scalable and can benefit from parallel processing for the computing and data intensive tasks. This will allow the system to deal with varying workloads for large organizations and provide analysis response times [48] in line with police investigation requirements.

3. Tasks to be solved

The functional and requirement analysis of the methods will be worked out as well as the standardization task. This work should ensure that the developed tools meet the requirements of the end-users and the police forces, and it is therefore considered as a key work package.

For each of the considered multimedia resources, one considers the special needs related to the automated processing of the specific data. Therefore, five tasks are related to the treatment of images and videos, text, handwriting and audio and speech signals. Feature extraction methods are considered for all multimedia sources in work package. This strategy guarantees the synergy that can be obtained in quality improvement for processing when using information from different media types. The reasoning unit is considered in a single work package. All developed method will be linked to the CBR system. They provide a case description for each new case to the CBR system.

Novelty detection has aspects that go beyond CBR. Therefore, the development of the novelty detection unit is left to a single task.

The legal aspects are worked out in a special task. It should run over the whole project with the aim to ensure that each RTD task meets the legal aspects as well as to identify new aspects that arise when developing novel techniques for the analysis of forensic data.

All the proposed methods should be integrated into a single reasoning system. After this has been done, the final evaluation of the system will be done. The final system will be demonstrated to police forces and end users.

The proposed methodology of the work is shown in **Figure 1**.

3.1 Video and image enhancement, filtering and assessment

The goal of this package is to improve the quality of an image or video sequence in order to support easy classification or recognition tasks and to detect in images traces of tampering without using protecting pre-extracted or pre-embedded information. Fragile watermarking schemes to protect legal evidence audio and speech data will also be developed.

In many image modalities, synthetic aperture radar (SAR) images, passive millimeter wave (PMMW) images, commercial photography and images and videos acquired for security purposes, the captured images and video represent a degraded version of the original scene. The observed degraded image is usually the result of convolving the original image (to be estimated) with a known or unknown blurring function and the addition of noise. The removal of noise and blur from the observation in order to obtain a good estimate of the original image is the goal of deconvolution and blind deconvolution techniques. In this task, we will develop and utilize Bayesian deconvolution and blind deconvolution techniques to improve the quality of the observed images. The improved images will either be used by humans for identification or be the input to classification and recognition methods.

3.1.1 Image and video super resolution

We use the term super resolution (SR) to describe the process of obtaining a high-resolution image or a sequence of high-resolution images from a set of low-resolution (LR) observations. In this task we will utilize and develop motion-based spatial super resolution techniques as well as motion free super resolution techniques. In motion-based SR, the LR observed images are under-sampled, and they

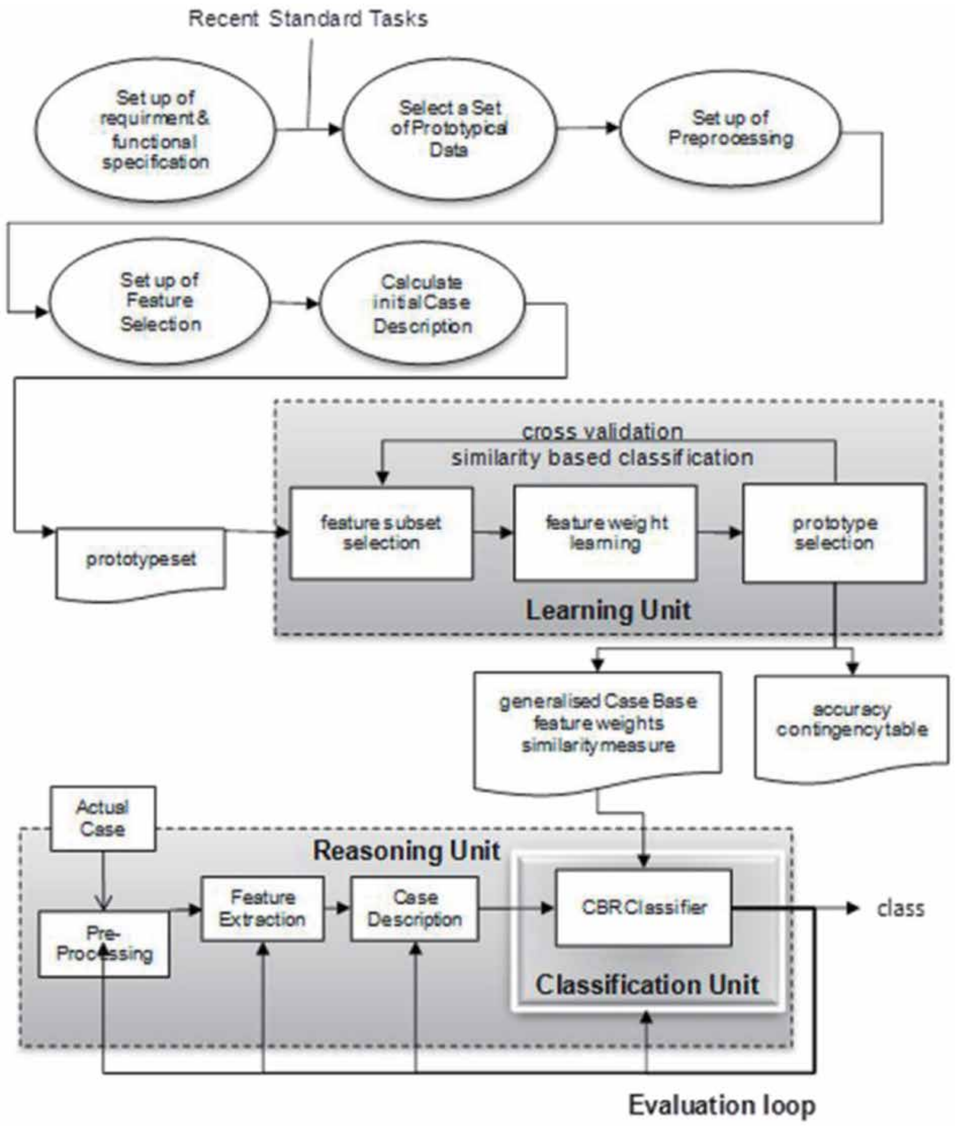


Figure 1.
Methodology.

are acquired either by multiple sensors imaging a single scene or by a single sensor imaging a scene over a period of time. In motion free SR images are upsampled by learning the relationship between corresponding low resolution and high-resolution image patches in a database and combining this learning process with the observed LR image. The improved images will either be used by humans for identification or be the input to classification and recognition methods.

3.1.2 Blind methods for detecting image forgery

In order to trust the information extracted from images and videos, it is necessary to make sure that the image and video have been recorded by a camera and that no artifact has been added or object removed. The trustworthiness of images and videos has clearly an essential role in many security areas, including forensic investigation, criminal investigation, surveillance systems, and intelligence services.

In this task we will utilize and develop blind methods for detecting image forgery, that is, methods that use the image function to perform the forgery detection task. The methods will try to identify various traces of tampering and detect them separately. The final decision about the forgery will be carried out by fusion of results of separate detectors.

3.2 Case-based reasoning

In this work package, we will develop novel case-based reasoning methods for a case-based reasoning system that can keep complex multimedia cases based on their different multimedia features and specific event features in a case base so that they can be easily retrieved and applied for new situations. Meta-learning methods based on CBR over the proposed multimedia processing chain will be developed in this WP in order to achieve the best processing results. The case-based reasoning system will consist of novel probabilistic and similarity-based methods. It will provide a wide range of novel similarity measures for the different feature types and representations for identification and similarity determination. Methods for hierarchical organization of the case base will allow very fast retrieval of similar cases. A special taxonomy for similarity determination and measures will be worked out and implemented in the CBR system. It will provide explanation capabilities for similarity, as those will help a forensic data analyst to identify the right reasoning method for his particular problem. This aspect goes along with the training and education aspect for forensic data analysis. Part of this will be self-contained in the chosen methods and realized by the system.

We will also develop learning methods to include new data into the existing cases and summarization of new and old cases into more general cases applicable to a wider range of tasks for further law purposes. The lifetime aspect of such a CBR system will be considered by special case base maintenance methods and modularity of the system architecture.

- Development of the system architecture

The main architecture of the CBR system will be developed, taking into account the different multimedia data types and data representations. The interface to the preprocessing units and the feature extraction units will be defined. The initial case description that can represent the different multimedia data types and data representations will be developed.

- Development and implementation of the case base for the different multimedia sources

The case base that is the heart of a CBR system will be developed and implemented. For the different multimedia-representation, the right database will be chosen as well as the right data structure. The interfaces and the data structure will be defined.

- Development and implementation of similarity measures for the different feature types and representation

An overview about different similarity measures for the different media type-representations will be developed. The pros and cons of the similarity measures will be worked out, and novel similarity measures for the respective data types will be developed. Aggregation of similarities of different types will be studied and evaluated.

- Development and implementation of a taxonomy over the similarity measures

As an outcome of these tasks, a taxonomy over similarities will be developed. This taxonomy will be represented by a hierarchical concept, and the user will be guided through this taxonomy for his special needs. A conversational strategy will be developed that helps the user to figure out what his needs are.

- Development and implementation of an indexing structure

The indexing structure over the different data types will be developed and implemented.

- Development and implementation of meta-learning methods over image and video processing chain

The preprocessing and feature extraction methods developed will be evaluated, and it will be decided where meta-learning has to be applied. The architecture and CBR methods for meta-learning will be defined. The meta-learning architecture that fits the CBR system will be developed. The meta-learning algorithm will be implemented and evaluated.

- Development and implementation of a generalization mechanism over cases, case-classes and higher-order classes

The methods for meta-learning over case, case-classes and higher-order classes will be defined for the different multimedia data sources. The interaction between these different multimedia-representations will be studied, and methods that can improve the performance will be developed.

- Development of a learning mechanism over similarities

A recent learning mechanism will be studied, and based on that a new learning mechanism will be developed for feature weighting of the different data types and deciding about the correct similarity measure. These strategies will be implemented into the CBR system and evaluated.

- Development and implementation of a life cycle and maintenance function for the case-based reasoning system

The life cycle of a multimedia CBR system will be studied based on the experience the partners have with different data sources. A life cycle and maintenance function that can take this into account will be developed and implemented.

3.3 Multimedia feature extraction

This work package will investigate, define and evaluate feature extraction methods to detect, describe and relate the multimedia data content relevant to forensic activities. The activities will concentrate on different biometric parameters that characterize individuals in terms of appearance, behavior, voice and handwritings, so as to enable the process of detection and recognition.

Features pertaining to face characteristics, distinctive individual marks, morphometric measurements of the body and gait analysis will be the subject of investigation for videos and images. Similarly, the exploration will regard distinctive features from audio signals for speaker's identification and recognition as well as text analysis to extract information pertinent to the forensic investigations.

Particular emphasis will be put on texts whose language shows characteristics deviating from the standard written form: this will be the case of transcriptions of speech recognizers as well as of the language of social media.

Aiming at recognition in the wild, focus will be given to the definition and verification of features that enable detection and recognition in unconstrained conditions and environments. This means that feature invariance to different condition's changes and robustness to noise will be two fundamental issues that will be tackled. A systematic approach will be used to feature organization. This means that, starting from existing metadata standards for image, video, audio as well as textual data, a formal ontological model will be defined to organize and categorize all the features collected from pertinent literature as well as the ones newly defined. The resulting feature ontology will standardize feature definition and computation, catalog features and model the multimedia data analysis domain. Such an ontology will be integrated with a library of algorithms for the computation of the features considered, resulting in a toolbox for feature extraction.

- Feature extraction from images and videos

This task will focus on the definition and extraction of features to be used in detection, recognition, authentication and tracking of individuals, event analysis and anomaly/novelty detection.

The work will concentrate on biometric and general appearance features. The field of biologically inspired features will be investigated as well, to verify whether methods that try to mimic the human visual capabilities are able to ensure a better performance. This will be particularly addressed to tackle the difficulties of forensic scenarios where it is not possible to make restrictive assumptions about ambient illumination, subject pose, sensor resolution and compression. In particular, for face recognition in the wild, three-dimensional features will also be explored.

The approach based on bags of features will be studied to carry out a first 'skimming off the top' in large repositories to find out only relevant objects pertinent to the case at hand. Local descriptors will be then defined to handle intensity, rotation, scale and affine variations. Improvements based on the inclusion of spatial information will be investigated.

- Feature extraction from audio streams

This task will focus on the definition and extraction of features for audio streams for the purpose of (i) identification/authentication of individuals and (ii) audio event detection.

The following two main lines of research will be followed: (i) Audio diarization where first the audio stream will be first segmented in speech, nonspeech audio (incl. music) and background noise. Then the speech portions of the audio streams will be segmented by speaker turn and identity. Finally the speaker identity will be verified or identified (depending on the scenario). (ii) Computation of the saliency of the audio stream using low-level features to identify surprising audio events. The audio events will be then classified into semantic (ontological) categories that will be used in Task 3.4.

The audio feature extraction package will include generic short-time envelope features (e.g., mel frequency cepstrum coefficients (MFCC), perceptual linear prediction coefficients (PLP), spectral envelope coefficients (SMAC)) as well as time-domain features. In addition, for speaker identification/verification, modulation spectrum and micro-modulation (AM-FM) features will be employed.

- Feature extraction from text

This task will focus on the definition and extraction of features from running texts carrying relevant content for forensic activities. This goal will be pursued by combining machine learning stochastic algorithms with advanced natural language processing techniques in line with the state of the art in the computational linguistics field. Two main lines of research can be envisaged for this task, respectively, aimed at (a) extracting ontological knowledge from texts to be used in the framework of Task 3.4 for building the feature ontology and (b) recognizing and semantically classifying relevant information in running texts, to be used as features describing individual documents. We will refer to these lines of research as 'ontology learning' and 'feature extraction', respectively. The planned work will mainly consist in the customization and integration of pre-existing software components available from the consortium partners contributing to this task, which will be specialized to meet the specific needs of particularly text. In particular, the main customizations will be concerned with the typology of information to be extracted (also including relational information) as well as with more challenging research topics such as the automatic analysis of texts representative of so-called noncanonical languages or the development of sophisticated technologies devoted to false witness detection. The final result of this task will include automatically extracted ontological knowledge (to be used as input to Task 3.4) as well as tools for feature extraction from texts to be possibly included in the toolbox for feature extraction as far as texts are concerned.

- Feature integration and ontology development

All the features developed and collected in the other tasks will be precisely defined and formalized following a coherent feature definition model. This will enable the development of an ontological model to accommodate the different classes of features and give them an easily sharable and reusable standard organization. The final aim is to (i) standardize and homogenize the feature terminology, (ii) collect and disseminate structured classes of features and (iii) support the choice and computation of features according to a method-oriented strategy. Moreover, a library of algorithms will be supplied and linked to such an ontology, resulting in a toolbox for feature extraction.

3.4 Text mining

The overall goal of this task is the design and development of methods and techniques for supporting human experts in the analysis of social media textual data. In particular, novel methods will be developed to (a) monitor in real time Twitter and identify potential threats including individuals and communities of users who are planning illegal activities and (b) build a dynamic model on Twitter text to forecast the upcoming significant events and emotions of the crowd associated with these events. Different approaches to the analysis of this type of texts will be pursued, including natural language processing (NLP) techniques, whose results will be eventually compared and—possibly—combined. The languages dealt with will include Dutch, Italian, Hungarian, English and Bulgarian.

- Linguistic analysis of Twitter data

In this phase the tweets will be linguistically analyzed: the text will be segmented in sentences, tokenized, morphologically analyzed and lemmatized. Linguistic preprocessing is needed to extract information from text: however, the linguistic analysis of small-sized texts, such as tweets, is nowadays a challenging task. It is a widely acknowledged fact that NLP systems, typically trained on newswire texts, have a drop of accuracy when tested against these kinds of texts. In tweets, punctuation and capitalization are often inconsistent, slang and technical jargon are widely exploited, and no-canonical syntactic structures frequently occur. This task is aimed at devising and testing domain adaptation methods to allow the NLP tools to achieve reliable results on these types of texts.

- Keyword extraction

In the first phase, keywords for early warning indicators on certain selected types of crimes will be developed. Traditional fully automated keyword extraction techniques have shown to perform poorly on small-sized texts. We will consider semi-automated keyword extraction methods. In particular, we will start from a domain specific set of keywords for football hooliganism provided by experienced police officers. This collection will unavoidably be incomplete; however, generic keywords in this set can help us zoom in on a subset of the entire dataset. Starting from this manually built set of keywords, we will develop methods to extract other relevant keywords. This will be done by exploring different strategies, also based on NLP techniques. The approaches that generate more reliable results will be used for continuously extending the set of keywords. Some of these methods will also be exploited in the feature extraction process from text carried out in the framework of WP3.

- Visualization

In the next phase, the Twitter feeds which remained after applying the early warning indicators should be visualized in an intuitive to interpret manner such that police officers can swiftly zoom in on potentially dangerous conversations. We plan to first use self-organizing maps which are built using a training algorithm that allows for incorporating user-defined and automatically inferred attribute priorities. The SOM will partition the collection of tweets into risk areas. The user can choose to group tweets based on the person who wrote them, and the map will show in this case the distribution of persons. If the user would like to analyze in detail a person, a Twitter conversation or a group of persons, we intend to provide functionality such that he or she can select an object or a collection of objects of interest and analyze it with a formal concept analysis and its temporal variant.

- Twitter data analysis

An important step is to identify communities of users in these large amounts of Twitter data. We hereby extend our analysis methods from object-attribute data to object-object data. Fuzzy or probabilistic measures depicting the strength of the relation between individual objects should be used to preprocess the data in order to cope with scalability issues. Based on these distance-related measures, the data is segmented, and strongly related subcommunities are extracted. These subcommunities can be investigated with traditional social network analysis methods, complemented by temporal concept analysis, temporal relational semantic systems, etc. in order to infer its threat level and the role of the actors in the network. To facilitate this phase, we need linguistic methods which will be used to extract relational attributes from Twitter data and can complement the keyword attributes.

- Development of text classification and regression techniques

Finally, automated classification and regression techniques should be developed to partially automate the suspect identification process. Initially a large amount of manual labor should be used to extract keywords, identify priorities among attributes, etc. In the later phases we aim at automating this process and the previously discussed visualization methods stay useful for explaining the automatically made decisions as well as validate them.

- Forecasting events relevant to legal forces

To provide more effective results, automatic identification of upcoming threatening events is essential. We will develop methods to identify events unknown beforehand, especially with potential interest to legal forces. We plan to build a dynamic model on Twitter text to forecast the upcoming significant events and emotions of the crowd associated with these events. While there can be many events with a strong presence in social media, some of them would have stronger negative emotions associated with them. These events are candidates that may have criminal nature or significant social consequences.

3.5 Video analysis

The huge amount of CCTV systems has increased the importance of video and image evidence in forensic labs. In order to retrieve the frames of interests, human experts usually spend a lot of time to examine hours and hours of video sequences. The low quality of the images due to high compression algorithm or bad light conditions or video coming from different record source multiplexed in a unique streaming makes the problem quite difficult to tackle. Another important aspect is the great number of native file formats of CCTV video and the difficulty to create a working copy of the data from a format conversion. This WP will develop novel automatic video processing tools aimed at supporting the expert in selecting the portions of videos that might contain the interesting facts he/she is looking for. In particular, techniques for people identification based on face recognition will be devised. As people may appear in nonfrontal poses, we will exploit the presence of multiple cameras in a given area to track a person and make the identification when the face is in frontal position.

- Video sequence analysis

To reconstruct the dynamic of the event of interest, the expert may require a lot of time spent to examine hours and hours of video sequences. A lot of factors can make the operation quite difficult in the absence of a suitable automatic software: low quality of the images due to high compression algorithm, bad light conditions or video coming from different record source multiplexed in an unique streaming. In addition, as multiple native file formats for CCTV video have been produced, the expert should also produce working copies of the data by normalizing the video format.

This task will be aimed to develop a tool that automatically separates video sequences coming from different cameras and converts the video sequence in a common format that can be used for further processing steps, e.g., feature extraction.

- Frame selection

This task aims at producing a semi-automatic tool to assist the human expert in selecting the most meaningful frames. Due to the huge quantity of manufacturers operating in the CCTV marketplace, there are a broad range of system for retrieve and export images from compressed video. This task will analyze the most common formats and produce a tool that will leverage on already available retrieval techniques provided either by the developer of CCTV system or by the open-source community.

- Person and object retrieval and identification

This task will be aimed to develop tools to aid the human expert looks for individuals or objects that are useful for the investigation. Object of interests can be, for example, heads, vehicles, license plates, guns, dresses and all other objects that can link a person to the event, etc. The tool will handle the cases of videos with low quality, bad lighting condition, camera/object position and facial expressions.

The persons/objects with enough quality retrieved from the video will be compared by an automatic procedure to a set of known elements of comparison. The expert will then provide feedback on the comparison results in order to refine the search.

An important main focus of police work is the identification of people for which a decision of the public prosecutor's office or a judge to the observation or an arrest warrant was issued. Within this scope of arrangement, the use of video-supervised places and facilities should be used. At earlier not known places, the application of mobile videotechnology should be deployed.

The aim of this task is to develop methods and procedures for an automatic system for identification of one or several target people in mobile video recordings based on passport photos or other available pictures. On this occasion, the prototype-based methods should be used, which are able to work on different picture representations, like pixel, features and graphs. By means of case-based learning mechanisms, a model should be automatically learnt for procedures to the facial recognition under the described application terms. In detail, the following should be developed:

1. Prototype-based methods and procedures for the identification of an aim person or personal group based on one or several prototypes obtained from pictures of the target person or target personal group without that the system on several picture sequences about likelihood must be trained.

2. Methods and procedures which allow it from a picture of the person (passport photo or photo of an observation) on site or fast to generate a prototype picture, which includes aging processes of the person and different perspectives and eliminates covers of clothes must be included without special facilities of the forensic disciplines.

3. To be able to learn methods and procedures to the generalization about case uses around the model for the facial recognition under these application terms.

- People reidentification

This task is aimed at devising techniques for the re-identification of people in videos captured by different cameras. In many outdoor and indoor environments, different cameras are present to monitor a given area (e.g., in a given street, you can find cameras operated by shops, banks, etc., or by the municipal police). The development of reidentification techniques allows tracking a subject that exits from the field of view of a camera and enters into the field of view of another camera placed in the neighborhood.

3.6 Speech and audio processing

A significant portion the data collected by law enforcement agencies are speech, sound and audio files.

Similarity, case-based reasoning, compressive reasoning and sensing-based novel methods for speech and audio recognition using both temporal and frequency domain information will be developed. Similarity learning, case generalization and case storage and compressive learning and sensing will allow the handling of very large amount (terabytes) of data.

- Speech and audio representation

Development of speech and audio representation schemes using both spectral, wavelet scattering and temporal methods such as delta modulation and zero-crossing information. This novel representation will be used in all of the above tasks.

- Case and similarity-based reasoning-based speech and audio recognition

Development of 'query by example,' keyword and phrase-based retrieval schemes using conventional and structural similarity-based methods capable of part and whole similarity matching. Once the keyword and phrases are detected, analysts can manually process the proposed retrieval results.

- Compressive reasoning and recognition

Exemplar-based speech, speaker and audio recognition. The resulting scheme will be computationally efficient, and it will work in the compressed data domain.

3.7 Handwriting recognition

Handwritings constitute another important part of the collected data. The objective of this work package is to develop computer vision and pattern recognition methods to identify and recognize the large volumes of unconstrained handwritten text to assist the experts.

Methods will be developed not only for writer identification which is a very important task in forensic applications but also for automatic recognition of the text found in the environment such as notes and letters written by the subject or the scene text in the photos taken, such as the shop labels and billboards.

Image enhancement techniques will be carefully applied to reduce the noise without deforming the original data. Alternative to character-based systems, word-based systems will be developed to read the text 'in the wild' using methods inspired from object detection and recognition literature. Generic image descriptors such as salient points, gradient histograms and line pairs will be considered for describing words. Efficient and effective similarity measures will be developed to match handwritten text in large volumes in a fast manner without losing any important data.

- Representation of words

Words in handwritten text will be described with advanced visual features used in generic object recognition. Novel descriptors based on contours, salient points and shapes will be generated.

- Similarity-based word matching

Development of 'query by example,' subword, keyword and phrase-based retrieval schemes using conventional and structural similarity-based methods capable of part and whole similarity matching.

- Writer identification

Documents will be preprocessed for enhancement and noise removal. Exemplar-based word and subword matching will be used to identify and verify the writers. Experts will be provided with a set of results and will be asked for feedback to be used in an active learning scheme.

- Automatic text recognition

Development of both character and word-based recognition schemes. Manual effort for labeling data during training will be reduced by learning the relationships between available handwritten and printed text pairs. Scene text will also be recognized.

3.8 Novelty detection

The objective is to develop a CBR-based novelty detection method based on a combination of statistical and similarity-based methods able to handle the different multimedia data types.

The method will perform novelty detection and handling of the novel cases for immediately reasoning and new model built up and should consider the incremental nature of the data. It will update the models incrementally based on MML and MDL methods.

- State of art on statistical novelty detection methods applicable to multimedia data types.

Different statistical model methods will be studied and evaluated for the different multimedia data types. The methods that are applicable to the data types will be selected.

- Model development for novelty detection classification for multimedia data types

Taking into account the outcome of Task 8.1, new methods for the different multimedia data types will be developed and implemented.

- Incremental model learning mechanism based on MDL or MML

The incremental learning mechanism for the updating of the models and feature description should be developed based on statistical learning methods such as MML or MDL.

- Incremental data collection mechanism and data base structure

Development of the data base structure for the different multimedia data types.

The incremental data collection mechanism will be developed, taking into account the data base structure.

- Task management mechanism between the model-based and the similarity-based unit

The task management mechanism that can handle the interaction and task division between the model-based module and the similarity-based module will be developed, implemented and evaluated.

- Integration into the CBR unit

The developed novelty detection unit will be integrated into the CBR unit and tested.

3.9 Legal aspects

To provide a clear picture of the current legal framework regulating the process of gathering, processing, analyzing and integrating multimedia data for security and judicial purposes at the European (EU Directives and Regulations) and national level.

To specialize the investigation on legal issues related to the use of sensible data for forensic purposes by the analysis of case studies and EU (ECJ) and national judgments.

To provide a framework of standards, quality indicators and approaches for the preservation and validity assessment of digital evidence for forensic purposes.

To support partners in sorting out potential ethical questions, by supporting the consulting process of the advisory body in solving questions specific to ethics, issues concerning sensible data processing and, generally, rules limiting the use of personal data extracted by massive data processing techniques:

- This task aims at providing a deep survey of the legal sources at the national and European level. The survey will outline the state of the art in EU law, namely, the first Data Protection Directive (Directive 95/46/EC), and will investigate the important role played by the European Court of Justice (ECJ) in its interpretation. The emphasis is on the principle of proportionality—the key concept in the ECJ's judgments—that requires that every specific instance of processing of personal data to be necessary for its concrete purpose. The so-called proportionality test has three components, which involve an assessment of a measure's suitability, necessity and proportionality strictu sensu. References to the test could be useful for evaluating the legal effects of the implementation of tools developed by the project. National legislation of the eight countries participating in the Consortium (FR, DE, IT, BG, GR, SP, TK and NL) will be collected, analyzed and compared, to provide a complete picture of the level of harmonization among European countries. The impact of 'proportionality test' in national judiciary will be tested.

- The second step makes specific reference to the new European rules under discussion (proposal for a Directive-COM-2012-2110 and Sec-2012-2072 final) and its accompanying documents (Impact Assessment by Commission staff working paper to the Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data and to the Directive 72). The implementation process of the new rules will be monitored, and their impact (especially the interaction between regulation and Directive), will be analyzed. Actual changes of the regulative framework in case of the enactment of the new legislation during the project life will be explained to partners by training activities.

- **Evaluation of digital evidence:** this task aims at explaining the concept of legal validity in the light of digital evidence and at pointing out the criteria (e.g. authenticity) on which legal validity of evidences can be assessed; Task 2 will be implemented in two steps and with a strong connection with the WP 2, by reference to the classification of typologies of data, formats and specific features.

- **Analysis of technical and legal requirements** (e.g. ISO Guide 25) for acquiring, processing, storing and preserving digital evidences, based on the selection of guidelines and best practices of relevant judicial cases and of significant literature. (i.e., ENFSI community (http://www.enfsi.org/))

- A**nalysis of the communication process** between the judge, who is the final evaluator ('free conviction of the court') and the expert, who plays the role of a 'mediator', expected to provide the basic elements (time, location, authenticity, etc.) on which the evidence is grounded. The cognitive background, the communicative interaction, the semantic/terminological mapping and the reasoning processes will be modeled in an integrated picture; on this topics Task 9.2 will publish a report that aims at providing the scientific community with an update and original approach to the methodology for quality assessment of digital evidence.

3.10 Evaluation

To integrate and test the different components in an application that can easily be deployed in a computing infrastructure. This task aims to perform integration testing of the components and building an easily deployable application.

To develop case studies that will illustrate the features of the system. The case studies will illustrate the need for video and image enhancement, filtering and assessment, case-based reasoning, multimedia feature extraction, video analysis, handwriting recognition and novelty detection and illustrate how the system complies with legal requirements.

To evaluate the developed methods in the integrated software solution and to demonstrate the applicability of the solutions.

Risk mapping on territory, or based on the results of the scenarios, to determine certain potentially illegal behavior occurrences involving the activation of protection and security, both in terms of allocation of armed guards and installation of dedicated electronic equipment;

Support the decision on whether to carry out an intervention and which mode to adopt, on the basis of the evolution of a given behavioral scenario.

- **Identification** of a large class of—potentially or effectively—illegal behaviors.

- **Registration and selective search** of details from an offense scenario

  So the solution needs to be evaluated against real life situations.

- Integration and testing

This task will integrate the components produced by the different tasks. This will require integrating the video and image components, the case-based reasoning components, the multimedia feature extraction component, the social media text analytics component, the video analysis and face detection component, the speech and audio processing component, the handwriting recognition component and the novelty detection component. The system build will then undergo integration and testing to ensure that architectural functional and nonfunctional requirements are satisfied. The resulting build architecture will be deployable on a cloud computing infrastructure so that scalability requirements can be tested and shown to meet the requirements.

- Case applications

This task will develop several case studies that illustrate the use of the system on real but anonymized data from the law enforcement agencies and end-user companies. These case studies will illustrate the use of video and image enhancement filtering and assessment (PMMW images), case-based reasoning (data provided by all data providers), multimedia feature extraction (data provided by all data providers), video analysis, handwriting recognition, text analysis and novelty detection for forensic analysis of multimedia data (data provided by all data providers). The case studies will also show how the system complies with legal requirements. The case studies will be incrementally developed through the different project iterations.

4. Expected results

As the expected result, we will have the following:

- Novel multimedia data acquisition and automatic analysis methods for forensic multimedia data such as images and videos, text, handwriting, speech and audio signals and social media data

- A novel toolkit for the automatic and semi-automatic analysis and interpretation of forensic multimedia data comprised of image enhancement algorithm

- A forensic case basis in which cases and generalized cases are stored for fast retrieval and reasoning

- A learning unit and a novelty detection unit for dealing with novel and formerly unseen data and situations and detecting new task for the standard

- New standards and a new methodology for the analysis of forensic multimedia data.

The toolkit will be used as benchmark for testing instruments and rules involving technical skills, operators and courts.

The scenario: the achievement of an effective set of technical solutions and of (law-compliant) procedural standards is a valuable goal for the ‘global society’, where cultural interaction, de-territorialization of social behaviors and interdependency of phenomena (e.g. environment, health, immigration, crimes prevention, anti-terrorism fight, etc.) require law makers and courts to face the evolution of social phenomena and their legal effects within a pan-European harmonized area of justice. European judges tend to adjust the national legal framework by referring to common and shared principles, thus invoking constitutional rules as higher principles to which anchor case solutions. The problem of adopting standards of conduct in judicial procedures seems no longer to be a questionable issue but rather looks like an undeniable fact at the center of a vivid discussion within the international community of jurists, scholars and courts. In a field of growing relevance, like digital evidences in the courts, the project will provide practices of use (cases, tests) and guidelines.

Knowing the procedures, understanding the purposes, and assessing the results.

When touching mostly the sensitive area of 'public rights' like freedom, security and equality, citizens must be ensured that their fundamental rights are secured and that States' actions should be directed to their protection against crimes at the minimum cost of their freedom. In the field of data mining, the question became, besides the regulative aspects, a matter of ethic. In this field the question is not simply a matter of what type of data is collected and whether it is relevant but also how it is collected and by whom. The fact that securely de-identified data can be collected without consent provided there is a legitimate purpose is a clear argument. But, still law enforcement agency and courts should have to legitimize the purpose in a way that citizens can understand. The project will produce public available reports on the technologies while explaining their use and application by means of transparent guidelines.

The business-oriented benefits of this project take the form of new techniques and tools for the analysis of forensic multimedia data that can be marketed as tool boxes or single software solutions for the specific tasks described in the proposal. It will make a marked improvement on the solutions currently in use by the companies involved in the proposed work and will open new markets for those that are not currently involved in forensics. It is also foreseen to establish new enterprises that market and further develop the proposed software solutions in the high-technology field of security. A special marketing and services entity that will advertise the tools in the security field and among police forces will also be established.

End-users will benefit from the more effective analysis of forensic data based on the standards and methodology.

5. Conclusions

With this chapter, we finish our work on multimedia forensic data analysis, which was started in Part I of Forensic Multimedia Data Analysis [1].

Forensic investigations on multimedia evidence usually develop along four different steps: analysis, selection, evaluation and comparison. During the analysis step, technicians typically look at huge amounts of different multimedia data (e.g. hours of video or audio recordings, pages and pages of text, hundreds and hundreds of pictures) to reconstruct the dynamic of the event and collect any piece of relevant information. This step obviously requires a lot of time, and many factors can make it difficult, among which data heterogeneity, quality and quantity are the most relevant. Afterwards, during the selection step, technicians select and acquire the most meaningful pieces of information from the different multimedia data (e.g., frames from videos, audio fragments and documents). Then, in the evaluation step, they look for relevant elements in the selected data, which will be further investigated in the comparison step. They can select heads, vehicles, license plates, guns, sentences, sounds and all other elements that can link a person to the event. The main problems are the low quality of media data due to high compression, adverse environmental conditions (e.g., noise and bad lighting condition), camera/object position and facial expressions. Finally, during the comparison step, technicians place the extracted elements side by side with a known element of comparison. From the comparison of general and particular characteristics, the operators give a level of similarity. In forensic application the use of automatic pattern recognition system gives poor performance because of the high variability of data recording. On the other hand, human perception is a great pattern recognition system but is characterized by high subjectivity and unknown reproducibility and performance.

In this chapter, we propose to develop a toolkit of methods and instruments that will be able to support analysts along all these steps, strongly reducing human intervention. First of all, it will include instruments to process different kinds of media data and, possibly, correlate them. This will obviously reduce the time spent to find the correct instruments for processing the medium at hand. Furthermore, it comprises preprocessing tools that alleviate, by filtering and enhancement, the problem of low data quality. In particular, for image and video data, a great help will come from super-resolution methods that will maximize the information contained in low-resolution images or videos (e.g., foster the process of face reconstruction and recognition from blurred images). This feature will greatly support all the subsequent steps.

Semi-automatic tools will be included to assist human experts in selecting the most meaningful pieces of data. In particular methods for the selection of frames from videos, images from large databases, keywords from text documents and pieces from audio signals will be developed for a first skimming of huge amounts of data according to criteria specified by the users. To this end, a great advantage will come from organizing feature extraction methods, which will also allow users to relate different types of media and operate on them contemporarily, when possible.

For evaluation and comparison, a toolkit will comprise advanced (semi-)automated instruments for different media processing so as to allow person and object retrieval, identification and recognition, writer identification, automatic handwriting recognition, speech and speaker recognition. All these methods will address the problem of recognition in the wild, going strongly beyond the current state of the art. Using the toolkit will ensure the reproducibility of the analysis and foster operators' objectivity.

Finally, the case-based approach will ensure that the knowledge acquired during each investigation will be suitably summarized, generalized and stored so as to be profitably reused in other investigations.

Taken as a whole, the toolkit will dramatically reduce efforts spent by operators in tedious and time-consuming tasks, such as retrieving and selecting multimedia data, thus focusing them on much more important investigations. Currently, there is no other toolkit on the market that addresses the analysis of different types of media and has such a broad range of applications. Usually, different and not integrated solutions are developed to tackle specific problems on single-modality data.

A fundamental concern in forensic investigation is the legal validity of evidence. We will deeply survey the legal frameworks at the national and European level, thus obtaining a clear picture of the legal hurdles governing data extraction, integration and use. Criteria and rules to evaluate digital evidence will be investigated; standards for analysis, production and usability will be acquired and, if necessary, extended. The results produced by the toolkit will then be appropriate in different national and international courts.

- Validity of data as proofs of evidence, guaranteed by the evidence usability criteria

- Objectivity of data analysis, ensured by the use of mathematical methods well documented and explained

- Traceability of the methods applied, obtained by logging all the tools selected and applied to each type of media

- Reproducibility of the investigation process

Standardization will be particularly promoted by two main outcomes:

- The case base, which will foster the spreading of similar procedures and protocols, since successful solutions will be stored and efficiently reused to support novel cases

- The feature ontological model, which will enable reproducibility and shareability of the features that can be extracted from multimedia data in different scenarios

ICT solutions such as our proposed toolkit put a big value in forensic activities, since they enable analysts to obtain a sound identification, preservation, recovery and presentation of facts and opinions pertinent to an investigation. The awareness of this capability has been spreading in the last years, and several research initiatives and industries have been focusing on forensic informatics.

Author details

Petra Perner
Institute of Computer Vision and Applied Computer Sciences IBaI, Leipzig, Germany

*Address all correspondence to: pperner@ibai-institut.de

References

[1] Perner P. Novel Methods for Forensic Multimedia Data: Part I

[2] Wang A. The Shazam music recognition service. Communications of the ACM. 2006;**49**(8):44-48

[3] Morris RN. Forensic Handwriting Identification: Fundamental Concepts and Principles. San Diego, USA: Academic Press; 2000

[4] Franke K, Schomaker L, Veenhuis C, Taubenheim C, Guyon I, Vuurpijl L, et al. WANDA: A generic framework applied in forensic handwriting analysis and writer identification. In: Proceedings of 3rd International Conference on Hybrid Intelligent Systems. 2003. pp. 927-938

[5] Srihari SN, Leedham G. A survey of computer methods in forensic document examination. In: Proceedings of the 11th Conference of the International Graphonomics Society. 2003. p. 279

[6] Schomaker L. Advances in writer identification and verification. In: Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Vol. 2. Parana, Brazil: IEEE; 2007. pp. 1268-1273

[7] Franke K, Koeppen MK. A framework for document pre-processing in forensic handwriting analysis. In: IWFHR00; 2000. pp. 73-81

[8] Srihari SN, Zhang B, Tomai C, Lee S, Shi Z, Shin Y-C. A system for handwriting matching and recognition. In: Proceedings of Symposium on Document Image Understanding Technology. 2003. pp. 67-75

[9] Niels R, Vuurpijl L. Using dynamic time warping for intuitive handwriting recognition. In: Proceedings of IGS2005. 2005. pp. 217-221

[10] Niels R, Vuurpijl L, Schomaker L. Automatic allograph matching in forensic writer identi-fication. International Journal of Pattern Recognition and Artificial Intelligence. 2007;**21**(01):61-81

[11] Pervouchine V, Leedham G. Extraction and analysis of forensic document examiner features used for writer identification. Pattern Recognition. 2007;**40**(3):1004-1013

[12] Madhvanath S, Govindaraju V. The role of holistic paradigms in handwritten word recog-nition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001;**23**(2): 149-164

[13] Rath TM, Manmatha R. Word spotting for historical documents. Journal on Document Analysis and Recognition. 2007;**9**:139-152

[14] Rodriguez-Serrano JA, Perronnin F. Handwritten word-spotting using hidden Markov models and universal vocabularies. Pattern Recognition. 2009; **42**(9):2106-2116

[15] Rothfeder J, Manmatha R, Rath T. Aligning transcripts to automatically segmented hand-written manuscripts. In: Proceedings of the Conference on Document Analysis Systems. Vol. 3872. 2006. pp. 84-95

[16] Leydier Y, Lebourgeois F, Emptoz H. Text search for medieval manuscript images. Pattern Recognition. 2007;**40**: 3552-3567

[17] Wang K, Belongie S. Word spotting in the wild. In: European Conference on Computer Vision (ECCV), Heraklion, Crete. 2010

[18] Plamondon R, Srihari SN. On-line and off-line handwriting recognition: A comprehensive survey. IEEE

Transactions on Pattern Analysis and Machine Intelligence. 2000;**2**(1):63-84

[19] Adamek T, O'Connor NE, Smeaton AF. Word matching using single closed contours for indexing handwritten historical documents [Internet]. Journal on Document Analysis and Recognition. 2007;**9**:153-165

[20] Ferrari V, Fevrier L, Jurie F, Schmid C. Groups of adjacent contour segments for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008;**30**(1):36-51

[21] Alahi A, Ortiz R, Vandergheynst P. FREAK: Fast retina keypoint. In: CVPR. 2012. pp. 510-517

[22] Ray H, Ess A, Tuytelaars T, Van Gool L. Speeded up robust features (SURF). Computer Vision and Image Understanding. 2008;**110**(3):346-359

[23] Ataer E, Duygulu P. Retrieval of ottoman documents. In: Proceedings of 8th ACM Internat. Work-shop on Multimedia Information Retrieval. 2006. pp. 155-162

[24] Ataer E, Duygulu P. Matching ottoman words: An image retrieval approach to historical document indexing. In: Proc. 6th ACM Internat. Conf. on Image and Video Retrieval. 2007. pp. 341-347

[25] Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. 2004;**60**(2):91-110

[26] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape con-texts. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;**24**(2):509-522

[27] Govindaraju V, Cao H, Bhardwaj A. Handwritten document retrieval strategies. In: AND '09: Proc. Third Workshop on Analytics for Noisy Unstructured Text Data, Barcelona, Spain. 2009. pp. 3-7

[28] Şaykol E, Güdükbay U, Ulusoy Ö. A database model for querying visual surveillance by integrating semantic and low-level features. In: Candan KS, Celentano A, editors. Proc. of 11th International Workshop on Multimedia Information Systems (MIS'05). LNCS. Vol. 3665. Heidelberg: Springer; 2005. pp. 163-176

[29] Tax DMJ, Jusycyak P. Kernel whitening for one-class classification. International Journal of Pattern Recognition and Artificial Intelligence. 2003;**17**(3):430-445

[30] Markow M, Singh S. Novelty detection: A review-Part 1: Statistical approaches. Signal Processing. 2003; **83**(12):2481-2497

[31] Scholkopf B, Platt JC, Shawe-Taylor J, Smola AJ. Estimating the support of a highdimensional distribution. Neural Computation. 2001;**13**:1443-1471

[32] Markou M, Singh S. A neural network-based novelty detector for image sequence analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;**28**(10): 1664-1677

[33] Weinshall D, Zweig A, Hermansky H, Kombrink S, Ohl FW, Anemuller J, et al. Beyond novelty detection: Incongruent events, when general and specific classifiers disagree. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2012;**34**(10): 1886-1901

[34] Perner P. Concepts for novelty detection and handling based on a case-based reasoning processing schema. Engineering Applications of Artificial Intelligence. 2009;**22**(1):86-91

[35] Şaykol E, Güdükbay U, Ulusoy Ö. Scenario-based query processing for

video-surveillance archives. Engineering Applications of Artificial Intelligence. 2010;**23**(3):331-345

[36] Şaykol E, Baştan M, Güdükbay U, Ulusoy Ö. Keyframe labeling technique for surveillance event classification. Optical Engineering. 2010;**49**(11): Article no. 117203, 12 p

[37] Marsland S. Density level detection is classification. Neural Computing Surveys. 2002;**3**:1-39

[38] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM Computing Surveys. 2009;**41**(3):1-15

[39] Can EF, Duygulu P. A line based representation to match the words in historical manuscripts. Pattern Recognition Letters. 2011;**32**(8):1081-1222

[40] Markow M, Singh S. Novelty detection: A review-Part 2: Neural network based approaches. Signal Processing. 2003;**83**(12):2499-2521

[41] Zhang Y, Callan J, Minka T. Novelty and redundancy detection in adaptive filtering. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 2002. pp. 81-88

[42] Zhang J, Ghahramani Z, Yang Y. A probabilistic model for online document clustering with application to novelty detection. In: Advances in Neural Information Processing Systems 17, NIPS 2004. 2004

[43] European Legal Sources. Available from: http://ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm

[44] Jones KJ, Bejtlich R, Rose CW, Farmer D, Venema W. The Computer Forenisics Library: File System, Forensic Analysis, Real Digital Forensics, Forensic Discovery. Addison Wesley; 2005

[45] Casey E. Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet. Waltham, MA: Academic Press; 2011

[46] Jones KJ, Bejtlich R, Rose CW. Real Digital Forensics. Pearson Education; 2005

[47] Rochwerger B, Breitgand D, Epstein A, Hadas D, Loy I, Nagin K, et al. Reservoir—When one cloud is not enough. IEEE Computer. 2011;**44**(3): 44-51

[48] Yang H-C, Dasdan A, Hsiao R-L, Parker DS. Map-reduce-merge: simplified relational data processing on large clusters. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data (SIGMOD '07). New York, NY, USA: ACM; 2007. pp. 1029-1040. DOI: 10.1145/1247480.1247602

Chapter 9

Explainable Artificial Intelligence for Digital Forensics: Opportunities, Challenges and a Drug Testing Case Study

Louise Kelly, Swati Sachan, Lei Ni, Fatima Almaghrabi, Richard Allmendinger and Yu-Wang Chen

Abstract

Forensic analysis is typically a complex and time-consuming process requiring forensic investigators to collect and analyse different pieces of evidence to arrive at a solid recommendation. Our interest lies in forensic drug testing, where evidence comprises a multitude of experimentally obtained data from samples (e.g. hair or nails), occasionally combined with questionnaire data, with a goal of quantifying the likelihood of drug use. The availability of intelligent data-driven technologies can support holistic decision-making in such scenarios, but this needs to be done in a transparent fashion (as opposed to using black-box models). To this end, this book chapter investigates the opportunities and challenges of developing interactive and eXplainable Artificial Intelligence (XAI) systems to support digital forensics and automate the decision-making process to enable fast and reliable generation of evidence for the court of law. Relevant XAI techniques and their applications in forensic testing, including feature section, missing data handling, XAI for multi-criteria and interactive learning, are discussed in detail. A case study on a forensic science company is used to demonstrate the real challenges of forensic reporting and potential for making use of forensic data to pave the way for future research towards XAI-driven digital forensics.

Keywords: digital forensics, drug testing, machine learning, explainable AI, decision-making, automation

1. Introduction

The primary focus of forensic analysis is the acquisition of accurate and reliable evidence through the utilisation of methodologies that have proven consistent and trustworthy across the domain [1]. The evidence is presented to the court of law and the prosecutor must be satisfied with its reliability, credibility and admissibility. Forensic evidence can be extremely sensitive and dangerous for law enforcement to handle and the use of incorrect or unreliable evidence threatens the safety of justice.

Digital forensics (DF) was introduced as a means of digitally making use of forensic data for both the discovery and interpretation of electronic evidence [2]. This area has become increasingly important with the surge in the volume, variety and velocity of forensic data. Currently, the major challenges faced by DF investigators are an increase in the number of cases and the complexity of cases [1]. The increase in cases could be due to a societal shift towards faith in DF techniques, with the common belief being that advanced tools are highly useful in skilfully extracting and using forensic information [2]. The increasing complexity of cases is simply a result of advances in technology, storage and applications [1]. Another challenge for DF investigators is the requirement for fast turnaround. Due to the nature of forensic inquiries, investigators wish to have faster, more advanced and more accurate tools, in order to prevent any setbacks that could adversely affect the case. Furthermore, it is expected that new challenges will arise for DF in the near future as pointed out by Mazurczyk et al. [3], p. 10: 'modern digital forensics is a multidisciplinary effort that embraces several fields, including law, computer science, finance, networking, data mining and criminal justice'.

Artificial Intelligence (AI) is a technology that has been used for many decades, with growing importance in the modern day due to its uses for learning and reasoning. AI methods are extremely capable of learning and solving complex computational problems and have subsequently been considered crucial for future developments; from explaining the reasoning process of expert systems, to recognising patterns in artificial neural networks [4, 5]. Although AI models have been developed to support parts of the court cases, current judiciary systems may raise concerns over the reliability of decisions made by AI models. Moreover, these models can be useful but only when explained to judges and jurors, such as in a study by Vlek et al. [6] where they used scenario scheme idioms to construct Bayesian Networks (BN), in order to make the network easier to understand. This method attempted to explain why certain modelling choices were made as well as why the network arrived at the final output, given the choices made along the way. Another paper by Timmer et al. [7] used BNs to formalise the relationship between the hypothesis and the evidence presented in the network, and the authors derived a support graph to assist with interpretation of the BN, which could then be used for argument and evidence about the case.

In view of the importance of explainability, there emerges XAI, a collection of AI methods that focuses on producing outputs and recommendations that can be understood and interpreted by human experts. A focus of the AI community at the moment is to develop XAI methods that have a good balance between both transparency and explainability as well as power, performance and accuracy [8]. The application of XAI models to DF problems is scarce but would open up the possibility of using computer-based analysis for evidence in courts of law. It could become an extremely powerful tool for helping judges and jurors make decisions in the presence of many interconnected pieces of evidence.

This chapter investigates the opportunities and challenges of applying XAI to support DF. First, this chapter discusses DF and the applications of AI in the forensics domain. Second, it reviews existing literature on XAI, feature selection methods built on various types of variables such as images and electrodermal activity for drug and alcohol testing, missing data handling techniques and XAI for multi-criteria and interactive learning and their implementation in DF. Third, it discusses a current case study on drug testing that includes problem formulation, a description of the forensics data collected from questionnaires and analytical testing, and the high-level decision-making process for drug screening. Finally, the chapter presents a conclusion drawn from this study and further work.

2. Background

This section puts this chapter in context by reviewing the area of XAI and its application to DF, and discussing several data-related challenges one may need to address to make the most out of XAI methods, such as dealing with a large number of variables/features, missing data, multiple (conflicting) output criteria and interactions between the AI system and the practitioner.

2.1 XAI and its application in digital forensics

With ML being the core technology, AI systems have made remarkable achievements in solving increasingly complex computational tasks and making them critical aspects of the future development of human society [4]. However in case of ML algorithmic models pursuing prediction accuracy and becoming increasingly opaque, the explainability becomes problematic for black-box techniques such as ensemble methods and deep neural networks [9].

To address the trade-off between interpretability and model performance, post-hoc interpretability techniques emerge, which approximate black-box models by techniques such as simplification, feature relevance estimation, or visualisation. Eventually, the opaque models are turned into glass-box, which achieve a good trade-off between interpretability and prediction accuracy. Examples of such techniques include local interpretable model-agnostic explanations (LIME) [10], which explain the predictions by approximating the opaque black-box model with simple models locally, and SHapley Additive exPlanations (SHAP) [11], which calculate the contribution of each feature to the prediction based on three desirable properties (i.e. local accuracy, missingness and consistency). These techniques are referred to as XAI, which propose creating a collection of ML techniques that generate more explainable, understandable and trustworthy models without losing out significantly in prediction accuracy [8]. XAI methods can be classified according to multiple criteria, including intrinsic or post hoc, model-specific or model-agnostic and local or global interpretability [12].

2.1.1 Intrinsic or post hoc?

This criterion distinguishes whether XAI is achieved intrinsically or post hoc. Intrinsic interpretability refers to ML models that are interpretable because of their simple structures (e.g. linear models, tree-based models). Post hoc interpretability refers to the use of methods like feature importance and partial dependency plots in explaining the black-box models (e.g. ensemble methods, neural network) after training.

2.1.2 Model-specific or model-agnostic?

For model-specific techniques, interpretability is incorporated within the internals (i.e. inherent structure and learning mechanisms) and is limited to specific models. In contrast, model-agnostic methods, as named, are irrelevant to the inner processing/structure of the model. They can be seamlessly used on any ML model and are applied after the model has been trained [12].

2.1.3 Local or global?

The scope of the interpretability, global to the model or local to the prediction, is another important criterion [10]. Global interpretability refers to the entire model behaviour and answers 'show does the trained model make predictions?'. Local interpretation methods explain a single prediction which influences a user's confidence in the prediction and consequently, the user's action.

DF, which requires the intelligent analysis of large amounts of complex data, is benefiting from AI. Mitchell [5] reviewed some of the basic AI techniques that have been applied to the DF arena. These include expert systems in explaining the reasoning process, Artificial Neural Network (ANN) in pattern recognition, and decision trees acting as learning the rules for pattern classification and expert system. Irons and Lallie [13] also identified the use of AI techniques to automate aspects like identification, gathering, preservation and analysis of evidence in DF process. In recent years, the importance and requirement of using explainable methods which achieve both the robustness of algorithms and transparency of reasoning have been increasingly acknowledged in DF. Interpretable ML classifiers like decision trees and rule-based models have been commonly applied to DF problem [14, 15]. To explain a legal case, the community has also applied the idea of BN [6, 7]. AfzaliSeresht et al. [16] presented an XAI model in which event-based rules are created to generate stories for detecting patterns in security event logs for assisting forensic investigators. Mahajan et al. [17] applied LIME towards toxic comment classification in cyber forensics and achieved both high accuracy and interpretability compared to various ML models. However, in terms of automated decision-making in DF, there are very few works that have been made to make it explainable. **Figure 1** provides the classification of XAI techniques and their recent applications in DF.

2.2 Feature selection and dimensionality reduction

The increase in the availability of data due to a push in digitisation has led to high-dimensional data sets for training and testing AI algorithms. However, the

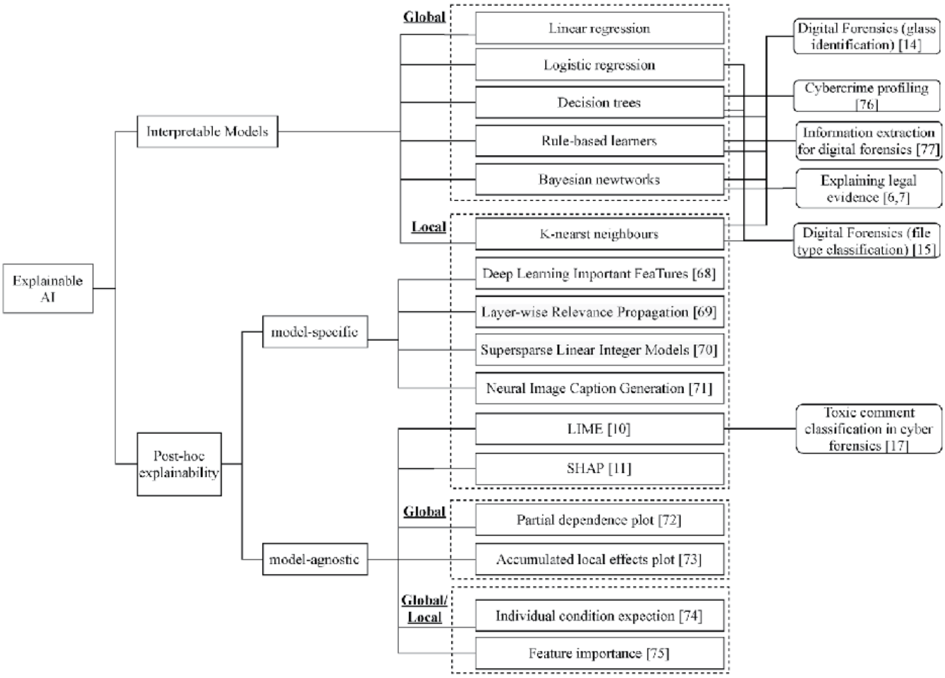

Explainable AI
Interpretable Models
Post-hoc explainability
model-specific
model-agnostic
Global
Linear regression
Logistic regression
Decision trees
Rule-based learners
Bayesian newtworks
Local
K-nearst neighbours
Deep Learning Important FeaTures [68]
Layer-wise Relevance Propagation [69]
Supersparse Linear Integer Models [70]
Neural Image Caption Generation [71]
LIME [10]
SHAP [11]
Global
Partial dependence plot [72]
Accumulated local effects plot [73]
Global/ Local
Individual condition expection [74]
Feature importance [75]
Digital Forensics (glass identification) [14]
Cybercrime profiling [76]
Information extraction for digital forensics [77]
Explaining legal evidence [6,7]
Digital Forensics (file type classification) [15]
Toxic comment classification in cyber forensics [17]

Figure 1.
Classification of XAI techniques and selected applications in DF.

amount of available data is just as important as the quality of the data. To ensure high-quality data is being filtered out from redundant, irrelevant, or noisy data [18], one can apply feature selection. Selecting the most relevant features has been shown to increase prediction accuracy, since it simplifies the model [19] and removes redundant in features [20]. However, the situation of having too little data needs to be avoided where possible to reduce the risk of overfitting, which occurs when a function is too closely fit to a limited set of data points. It is worthwhile highlighting the difference between feature selection and dimensionality reduction: while both methods reduce the number of features in a dataset, feature selection is achieving this by simply selecting and excluding given features without changing them, dimensionality reduction transforms features into a lower dimension. Our focus is more on feature selection methods. However, commonly used dimensionality reduction methods include Principal Component Analysis (PCA), Random Projection, Partial Least Squares and Information Gain.

Feature selection methods are categorised in **Figure 2** according to their process of ranking features into filter, wrapper and embedded techniques [21]. Filter methods are techniques that rank the relationship of features with an outcome without learning a model, such as Separability and Correlation Analysis (SEPCOR) [20]. Univariate filters calculate the ranking for each individual feature, while multivariate filters compute the ranking based on the correlation between the variables or between the variables and the outcome [22]. Wrapper techniques select features by comparing all the combinations of the included features before starting

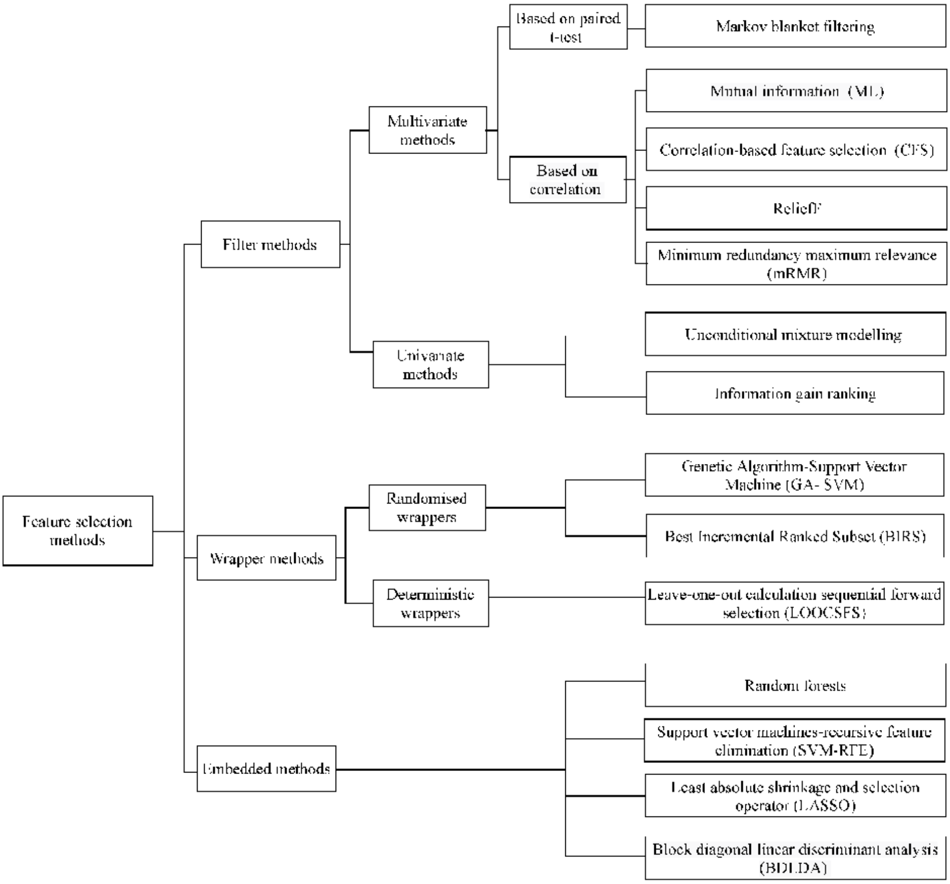

Feature selection methods
Filter methods
Multivariate methods
Based on paired t-test
Markov blanket filtering
Based on correlation
Mutual information (MI)
Correlation-based feature selection (CFS)
ReliefF
Minimum redundancy maximum relevance (mRMR)
Univariate methods
Unconditional mixture modelling
Information gain ranking
Wrapper methods
Randomised wrappers
Genetic Algorithm-Support Vector Machine (GA- SVM)
Best Incremental Ranked Subset (BIRS)
Deterministic wrappers
Leave-one-out calculation sequential forward selection (LOOCSFS)
Embedded methods
Random forests
Support vector machines-recursive feature elimination (SVM-RFE)
Least absolute shrinkage and selection operator (LASSO)
Block diagonal linear discriminant analysis (BDLDA)

Figure 2.
Classification of feature selection techniques.

the prediction model, to find the most accurately predictive one [22]. Wrapper techniques are more computationally expensive than filters; however, they generally produce more accurate results. Finally, embedded methods are classifier-dependent selection methods, where the selection is built based on the classifiers' chosen hypotheses [23].

Many comparative studies have been performed to find the best feature selection technique for high-dimensional data. For example, Hua et al. [24] compared a wide range of feature selection techniques for a variety of high-dimensional datasets. The authors followed a two-stage feature selection process to reduce computational time. In the first stage, feature selection methods that are independent from the classification process were applied. Following that, a further feature selection was implemented through classifier-specific feature selection techniques. The results show that wrapper methods have better performance in datasets with large samples, and filters have generally equal error trend. One of the main conclusions of their paper is that there is no feature selection technique performed best across all datasets. Another review of feature selection methods for high-dimensional datasets, which focused on filters, was conducted by Ferreira and Figueiredo [25]. The authors compared, amongst others, the following feature selection techniques for supervised learning: ReliefF, correlation-based filter selection, fast correlation-based filter, Fisher's ratio and minimum redundancy maximum relevance. Other solutions to tackle high-dimensionality in feature selection are the choice of an adequate evaluation criteria, such as predictive measures designed for small sample datasets and ensemble feature selection methods, including combining multiple feature selection methods and boosting [26].

Table 1 provides an overview of different feature selection methods applied to forensic science applications. Shri and Sriraam [20] formulated a feature extraction and feature selection problem to detect the difference between alcoholics and control groups through measuring the impact of the use of alcohol in multichannel EEG signal regions. Feature subset selection was performed using separability and correlation analysis, which was proposed in the paper. The results illustrate that the introduced technique improved prediction accuracy, and further validation using

Forensic application	Type of feature selection	Algorithm	Type of data	Reference
Alcohol testing	Filter method	Separability and correlation analysis	EEG signals, eye blink artefact and motion artefact	[20]
		Feature ranking using area under the curve	Continuous data	[27]
		Feature ranking using area under the curve	Categorical and continuous data	[28]
		Linear Discriminant Analysis (LDA)	Images	[29]
Screening substance use disorder		A discriminant function analysis	Categorical and continuous data	[30]
Drug testing		Linear Discriminant Analysis (LDA)	Mass spectral data	[31]
	Wrapper method	Exhaustive search method	Continuous and time domain features	[32]

Table 1.
Selected applications of feature selection techniques in forensic research.

other classifiers and cross-validation is recommended. Another feature selection technique to enhance screening of alcohol use disorder was introduced by Mumtaz et al. [27]. The EEG features were recorded in 5-minutes eyes open and 5-minutes eyes closed segments. The implemented feature selection takes two steps. First, the relevance of each feature to the outcome is calculated using the ROC. Then, Markov blanket filtering combined with the ROC is used to remove redundant features. The second step has a high computational cost, which is one of the drawbacks of this method. The paper found that the inter-hemispheric coherence between the brain regions ranked the highest in classifying alcohol use disorder (AUD). Mumtaz et al. [28] designed a rank-based feature selection technique in response to the high-dimensionality in the dataset. Feature ranking was computed based on the area under the curve of that feature and represented the relevance of the feature to the outcome. The minimum number of features was chosen by adding the features to the model sequentially, starting from the highest-ranked features.

Another alcohol use detection method based on thermal infrared facial images was examined in [29]. The dimensionality reduction was carried out using PCA combined with Linear Discriminant Analysis (LDA) [33]. It was shown that LDA worked well if the data had no missing data [34]. In an application for feature selection [30], applied discriminant function analysis for substance use disorder detection. This disorder is usually related to P3 amplitude,[1] addiction severity and impulsivity in predicting treatment completion. The research found that the P3 amplitude accounts for more variance compared to other variables.

Mahmud et al. [32] designed a method for quick detection of opioid intake using wrist-worn biosensor-generated data. The exhaustive search method was applied to seek a set of variables that achieved the highest accuracy. It helped to minimise the computational time and increased the prediction accuracy and sensitivity. Feature selection methods have also been applied to identify illegal drugs [31]. PCA followed by LDA was implemented for drug isomer differentiation. Three feature selection models that were tested included the full spectrum, exclusion of selected masses and the selected region, where ions are expected to contribute to the isomeric difference.

To summarise, feature selection methods have been implemented in forensic research and particularly for the detection of substance use. Their application covers various types of data, including images, EEG signals and time-series. Most of the reviewed methods were based on a filters approach. However, since most of these applications have selected the features for classification purposes, embedded techniques are designed to integrate the selection in the classification process. Therefore, it is important to investigate other embedded and wrapper feature selection methods.

2.3 Missing data

Forensic data contains a large number of features. A proportion of information in these features could be missing, which would reflect a different level of uncertainty because they are measured independently in laboratories [35]. High-dimensional forensic data presents challenges in establishing unbiased estimation and inference of ML models. Missing and uncertain forensic data must be treated in the data preprocessing stage, before the development of ML models. The deletion of incomplete instances and imputation of missing data is the most frequently used

[1] The P3 is a positive deflection of EEG that occurs when a low probability novel, target, or oddball stimulus is presented within a sequence of high probability non-targets or standards [30].

method of handling missing data, however the removal of the incomplete instances results in biased inference due to poor representation of complete samples [36, 37].

Statistical methods based on data imputation are largely utilised to handle missing data. The basic idea is to replace the missing values with the predicted values obtained from the observed data. There are three types of missing data—missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) [38]. The missingness mechanism by MCAR is independent of observed and unobserved data whereas, MAR is independent of unobserved data and dependent on the observed data. The missingness mechanism by MNAR is only dependent on unobserved data. The forensic datasets are usually MCAR type.

The missing forensics data can be imputed by methods such as Multivariate imputation by Chained Equations (MICE), Maximum likelihood estimation (MLE), Random Forest (RF), K-nearest neighbour (KNN) and MICE by Regularised regression. MICE run a series of regression models whereby each variable with missing data is modelled conditional upon the other variables in the data [39]. This implies that each variable can be modelled according to its distribution. The missing data can be imputed by MLE using the expectation-maximisation (EM) algorithm [40]. It iteratively solves complete data problems and then intuitively fills the missing data with the best guess under the current estimate of the unknown parameters in E-STEP, then re-estimates the parameters from the observed and filled-in missing data in M-STEP.

The method based on the RF called missForest was presented to impute missing continuous and categorical attributes [41]. It averages the multiple imputed unpruned classification or regression trees and estimates the imputation error by built-in out-of-bag error estimates of RF. A study showed that RF imputation method has less bias estimate and narrower confidence interval compared to MICE [42].

KNN imputes the closest instance in a multi-dimensional space by K-nearest neighbour imputation method. The similarity between two instances is measured by distance function such as Euclidean distance function. KNN imputation can handle instances with multiple missing variables without a need for the creation of a separate predictive model for each variable [43].

However, it suffers from the curse of dimensionality and could be computationally expensive as it searches for similar instances in the entire dataset.

A regularised regression model minimises the loss function by imposing some penalties. The superiority of regularised regression in terms of biases in imputed missing values in high-dimensional data is presented in [44]. In MICE by regularised regression the initial missing data are imputed by a simple method such as mean or frequency. The new parameters are estimated in the next iteration through the regression model and then missing values are replaced by predicted values. These steps are repeated for each variable with missing values. This procedure is conducted iteratively until convergence. After convergence, the final imputed data is utilised as input in a ML model.

2.4 XAI for multi-criteria problems

XAI techniques have shown promise in solving complex problems with multiple criteria. For example, decision trees, with tree-like structure in which internal nodes stand for tests on features and leaf nodes represent a class label [45], have been used as interpretable supervised classifiers in handling multi-criteria problems like medical diagnosis [46]. Vuong et al. [47] applied decision trees in forensic investigation to automatically produce detection rules used by the robotic vehicle in cybersecurity based on both cyber criteria (network, CPU, disk data) and physical features (speed, vibration, power consumption).

While decision trees can be adopted for visual reasons to highlight the most influential features in a classification process [48], rules have a textual description and are also readily seen in multi-criteria decision aiding [49]. The most common rules are IF-THEN which discretise a high-dimensional, multivariate feature space into a series of simple and explainable decision statements [50]. Karabiyik and Aggarwal [51] proposed an automated disk forensic investigation tool that leverages a dynamic knowledge base created using rules in the form of IF-THEN statements. Belief-rule-base (BRB), an extension of the IF-THEN rule base, has also been used to address multi-criteria problems [52, 53]. The inference of BRB system is explained by using the evidential reasoning (ER) approach [54], which allows the representation of both qualitative and quantitative data by using belief distributions and the aggregation of belief-based information. In addition to interpretable models, model-agnostic XAI techniques such as using an extended Shapley Value [55] and augmentation-based surrogate model [56] have been adopted in the multi-criteria decision aiding models to further assist in explaining the result of these models to decision makers.

XAI techniques have also been used to solve decision problems with multiple objectives. For example, Pessach et al. [57] proposed a comprehensive analytical framework based on the Variable-Order Bayesian Network (VOBN) model to support HR recruiters in global recruitment scheme in balancing multiple organisational objectives. Other XAI techniques/systems developed to solve multi-objective problems include V2f-MORL (vector value function based multi-objective deep reinforcement learning) [58] and fuzzy rule-based systems with multi-objective evolutionary algorithms [59].

Indeed, the goal of XAI techniques is to have the simplest rules which are understandable for humans without sacrificing the performance, although simplicity and performance are often conflicting objectives [60]. To achieve both accuracy and comprehensibility, the two important but conflicting classifier properties, Piltaver et al. [61] proposed multi-objective learning of hybrid classifiers (MOLHC) algorithm in which the sub-trees in the initial classification decision tree are replaced with black-box classifiers so that the complete Pareto set of solutions (a set of solutions that do not dominate each other but are superior to the remaining solutions in the search space) is more likely to be found. Similarly, with objectives of maximising the model ability while minimising the complexity, Evans et al. [60] used multi-objective genetic programming, another tree-based construction method in which trees are evolved from a population of candidates rather than constructed greedily in a top-down manner, to construct model-agnostic representation of black-box estimators.

2.5 XAI in interactive learning

Interactive ML is an iterative process of learning that includes the interaction between humans and ML methods [62]. It has been applied for multiple purposes, such as visual analytics [63], interactive model analysis [64] and event sequence analysis [65]. Jiang et al. [62] reviewed recent research in interactive ML and its application to solve a variety of tasks, discussed research challenges and suggested future work in the area. One of the recommendations for future work is to combine XAI with interactive ML. For example, complex ML algorithms can be simplified by using easy to understand algorithms, which helps the process of model building and parameter tuning.

Previous research combining XAI with interactive learning was done, for example, by Spinner et al. [63]. This research used XAI to explain the output of a ML algorithm, searches for limitations within the models and optimises them. In addition, global monitoring and steering mechanisms were applied. A user study with nine participants was included to test the system, and the results indicated positive feedback from the users. Many other applications of XAI for interactive ML were applied in the form of visual analytics. A modular visual analytics framework was developed for topic modelling, which allows users to compare, evaluate and optimise topic models using a visual analytics dashboard [66]. The design of the framework is interpretable by users and adjusts to their optimisation goal, which is based on time-budget, analysis goal, expertise and the noisiness of the document collection.

A review of visual interaction, supporting dimensionality reduction systems and covering interpretable models, was conducted by Sacha et al. [67]. The paper constructed seven possible scenarios for the application of interactive ML in dimensionality reduction. These scenarios included: interactive feature selection, dimensionality reduction parameter tuning, defining constraints and dimensionality reduction type selection. The paper found that some of the previous studies investigated a combination of these scenarios and the maximum number of combined scenarios in a paper was four. The paper also observed that some of these scenarios were studied more in the literature, such as the feature selection, data selection and parameter tuning scenarios. The application of XAI for interactive learning in forensic science has not been explored yet but it is easy to see that this approach can be beneficial in this domain; for example, where collection of evidence can be controlled (e.g. if is obtained experimentally) but is expensive and/or time-consuming, then a suitable approach may be to use XAI in an interactive fashion with a user, who can decide to terminate evidence collection prematurely upon retrieval of sufficient evidence.

3. Case study

This case study describes the process of forensic investigation by experts from an existing forensic science company. It will explore the challenges faced by forensic experts in making decisions based on factual and heuristic knowledge gained through years of experience. It will discuss the opportunity to utilise the forensic data to develop an interpretable and trustworthy system for automation of the decision-making process [68–77].

3.1 Reporting challenges faced by forensic experts

Currently, a trained expert in this company makes a decision based on a combination of factors, including the analysis of the testing sample and other, external factors such as chemical treatments and more. The expert then produces a report explaining the reasoning behind their decision, outlining different standards and classifying their decisions into one of a plurality of outcomes surrounding likelihood of drug use and exposure to drugs.

The decision regarding likelihood of drug use or exposure is based on a multitude of considerations, including the level of drug detected, the specific metabolites, the client's self-declarations and many more factors. When the decision process and report writing is conducted by individual experts, there can be some variability in the final decisions and reports that are produced. One of the main reasons for this is the high volume of features to be taken into consideration, which may all have different levels of importance. Another is that with so many features, it is not possible to cover every potential case that may arise and therefore it is difficult to set specific guidelines for the experts to follow. There is also the potential for subjectivity of the expert when making the final decision—an issue which is difficult to eradicate when relying on human judgement. This can result in disagreement amongst individual experts, or uncertainty where experts may find it difficult to draw conclusions based on the evidence provided. Such differences in subjectivity could be due to personal experience, length of time in the role, previous encounters in different cases and many other potential effects.

When a metabolite is detected the machine generates information on the amount that was present in the sample or, in other words, the level. This is a continuous value which can be used by the experts to make decisions on whether the client was using a particular substance, whether they were exposed or if the client has not been in contact with a drug at all. The levels at which the expert defines use or exposure are up for debate. It can be difficult for them to pinpoint exact values where the judgement tips from likely exposure to likely use, and further problems arise when considering different levels within each category (e.g. highly likely, likely, etc.). Without set levels experts are using their own judgement to decide which category the client falls into, which again leaves room for disagreement across the board.

The most significant problem from a business-efficiency point of view is the length of time that it takes to write a report. A significant increase of new report instructions has resulted in the need for automation, as the current personnel are under high levels of pressure and demand for quick turnaround.

The need for automation is therefore not only to improve accuracy and reliability, but also to speed up delivery times and free up the time of the experts to allow them to undertake other key responsibilities such as research, training and dealing with abnormal cases. The current problem requires a system for automatic decision-making and report writing for the outcome of drug testing, to produce reports suitable for presentation in legal cases.

3.2 Forensic data

The features in the forensic data are collected through a combination of questionnaire data—which is completed by the client being tested—and the outcome of tests using forensic laboratory equipment. Each row represents an individual case and each column represents a feature. The forensic investigator collects the essential evidence such as hair and nails, as well as carrying out a structured questionnaire. The questionnaire consists of a number of sections, with a combination of multiple-choice options and Likert scale questions. The document collects information about medical history, drug and alcohol use, hair and nail care.

Hair and nail samples are an easy, non-invasive way of collecting the evidence required to detect the chemical and biological substances, which identify substance use or exposure. Depending on hair growth and the length of strands this can show up to 1 year of drug history, although typically only a maximum of 6 months is used during testing. Body hair is taken if there is less than 1 cm of hair available on the scalp. A nail sample is taken if scalp and body hair are both unavailable and can show up to 3–6 month of drug history. The evidence from hair and nail samples may fail the forensic test (false-negative results) if a suspect repeatedly cuts hair and nails, or uses certain chemical treatments. The forensic data from the questionnaire could gather missing features when some of the follow-up questions do not apply to a client. For example, follow-up questions for pregnancy would only apply only to females. The data could also be subject to inconsistencies due to inaccurate or false self-reporting. This could be due to inability to remember and answer the questions. Drug and alcohol intoxication can inhibit memory alone, making it difficult to obtain accurate information on both the quantity of the substance used/exposed to, and the number of days use/exposure, as the client is asked to recall over a period of up to 12 months. The analytical data collected through forensic laboratory tests could also be missing if the metabolites are not present in the client's body, as this would mean further testing is not required. The reason for this is that the testing equipment looks for every possible substance in the sample, rather than selecting those that have been instructed for analysis. The false-positive and false-negative test results affect the data quality. It could be due to external contamination in hair and nail samples, or having little to no body hair.

This type of forensic data can be used to develop decision support tools to fully automate the decision-making process and validation of the experts' assessments against empirical data. The XAI model supports complex decision-making and can process large amount of data in minutes. The steps for the development of automated decision-making system in the forensic investigation are shown in **Figure 3**, where the relevant techniques are described in detail in Section 2 of this chapter.

3.3 Decision-making process for testing Drug X

The decision-making process for testing Drug X[2] follows a hierarchical structure with binary outcomes, which has been simplified into a small decision tree shown in **Figure 4**. The specific metabolites have been anonymised, instead these have been renamed as 'Metabolite 1', 'Metabolite 2' and 'Metabolite 3'. It is a snapshot of an interactive-decision-tree that allows visualisation and assessment of the entire decision-making process followed by an expert when drawing conclusions on whether or not a client has used or been exposed to Drug X.

First, based on the questionnaire data the expert will check to see whether the client has declared any use of Drug X in the last 12 months. If this is true then use is confirmed and no further testing is needed. If use has not been declared, based on the analytical data which has been extracted from the hair or nail sample, the expert will consider whether the data shows detection of the Metabolite 1 compound. If Metabolite 1 is detected, further testing is required to determine the levels of Metabolite 1 present in different sections of the hair as this will inform the expert whether the client has used or been exposed to the drug.

If Metabolite 1 is not detected, the expert checks for Metabolite 2. If Metabolite 2 is detected then it is concluded that the client has been exposed, but if it is not

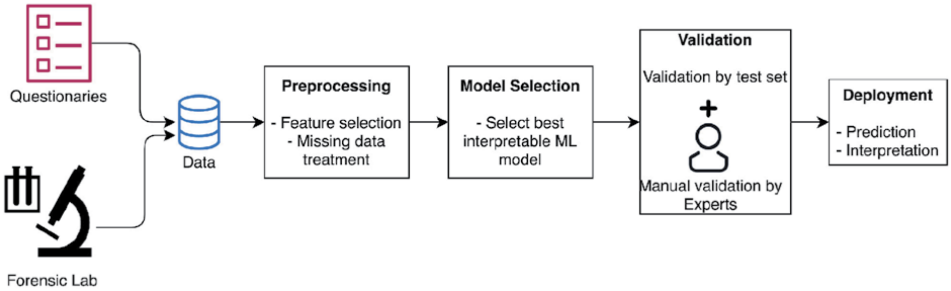


Figure 3.
Automated decision-making process.

[2] Drug X has been used to anonymise the name of the specific drug compound being discussed.

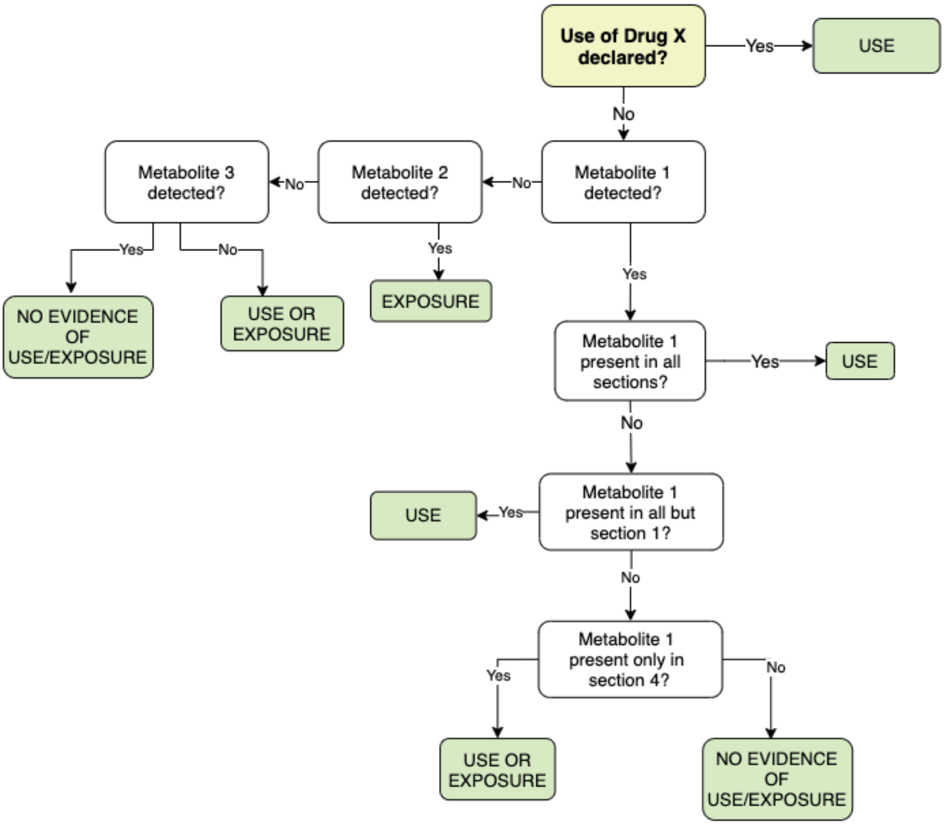

Use of Drug X declared?
Yes
USE
No
Metabolite 1 detected?
No
Metabolite 2 detected?
No
Metabolite 3 detected?
Yes
No
NO EVIDENCE OF USE/EXPOSURE
USE OR EXPOSURE
Yes
EXPOSURE
Yes
Metabolite 1 present in all sections?
Yes
USE
No
Metabolite 1 present in all but section 1?
Yes
USE
No
Metabolite 1 present only in section 4?
Yes
No
USE OR EXPOSURE
NO EVIDENCE OF USE/EXPOSURE

Figure 4.
Decision process for testing Drug X.

detected then the final check is for Metabolite 3. If Metabolite 3 is detected then it is determined that there is no evidence of use or exposure, but if it is detected then the decision is either use or exposure. This is dependent on the levels of each metabolite detected.

4. Conclusion and future work

This chapter has discussed the application of XAI to digital forensics with a particular focus on forensic drug testing. We provided an overview of data-related challenges one may face when implementing an XAI solution including a large number of features (e.g. pieces of evidence), missing data, multiple conflicting decision criteria and the need for interactive learning. Different techniques for dealing with these challenges were reviewed and applications in digital forensics were highlighted. Finally, we outlined a case study on a forensic science company to demonstrate real challenges of forensic reporting and the potential for XAI to design a trustworthy automated system to present generated evidence in the court of law.

The chapter proposes important future directions for adopting XAI techniques to address challenges in digital forensics. These include, first and foremost, the validation of the manually derived decision trees. It would be interesting to derive decision trees automatically using the available data. These trees could differ from the manually derived trees and thus reveal alternative drivers and potential hidden biases. Another direction is the development of more advanced XAI methods including belief or fuzzy rule based models. To make these data-driven models more accurate, one can also investigate systematic ways of merging with knowledge base and rules provided by experts. Thus, updating the rules can be done in an interactive fashion, for example as and when new scientific insight from chemistry becomes available. Certainly, these directions of future research are relevant for forensics in drug testing but also for digital forensics in general.

Author details

Louise Kelly*†, Swati Sachan†, Lei Ni†, Fatima Almaghrabi†, Richard Allmendinger and Yu-Wang Chen
University of Manchester, Manchester, UK

*Address all correspondence to: louise.kelly@manchester.ac.uk

† These authors are contributed equally.

References

[1] Golden G, Richard III, Roussev V. Next-generation digital forensics. Communications of the ACM. 2006; **49**(2):76-80

[2] Garfinkel SL. Digital forensics research: The next 10 years. Digital Investigation. 2010;**7**:S64-S73

[3] Mazurczyk W, Caviglione L, Wendzel S. Recent advancements in digital forensics. IEEE Security and Privacy. 2017;**15**(6):10-11

[4] West DM. The Future of Work: Robots, AI, and Automation. Washington, D.C: Brookings Institution Press; 2018

[5] Mitchell F. The use of artificial intelligence in digital forensics: An introduction. Digital Evidence and Electronic Signature Law Review. 2010; **7**:35

[6] Vlek CS, Prakken H, Renooij S, Verheij B. A method for explaining bayesian networks for legal evidence with scenarios. Artificial Intelligence and Law. 2016;**24**(3):285-324

[7] Timmer ST, Meyer J-JC, Prakken H, Renooij S, Verheij B. A two-phase method for extracting explanatory arguments from bayesian networks. International Journal of Approximate Reasoning. 2017;**80**:475-494

[8] Gunning D. Explainable Artificial Intelligence (xai), Web 2. Defense Advanced Research Projects Agency (DARPA); 2017

[9] Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020;**58**:82-115

[10] Ribeiro MT, Singh S, Guestrin C. “Why should I trust you?” explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: Association for Computing Machinery; 2016. pp. 1135-1144

[11] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. Red Hook: Curran Associates, Inc.; 2017. pp. 4765-4774

[12] Christoph Molnar. Interpretable Machine Learning. Lulu.com, 2019

[13] Irons A, Lallie HS. Digital forensics to intelligent forensics. Future Internet. 2014;**6**(3):584-596

[14] Tallón-Ballesteros AJ, Riquelme JC. Data mining methods applied to a digital forensics task for supervised machine learning. In: Computational Intelligence in Digital Forensics: Forensic Investigation and Applications. Switzerland: Springer; 2014. pp. 413-428

[15] Karampidis K, Kavallieratou E, Papadourakis G. Comparison of classification algorithms for file type detection a digital forensics perspective. Polibits. 2017;**56**:15-20

[16] Afzali Seresht N, Liu Q, Miao Y. An explainable intelligence model for security event analysis. In: Australasian Joint Conference on Artificial Intelligence. Switzerland: Springer; 2019. pp. 315-327

[17] Mahajan A, Shah D, Jafar G. Explainable AI approach towards toxic comment classification. In: Technical Report 2773, EasyChair. 2020

[18] Viegas F, Rocha L, Gonçalves M, Mourão F, Sá G, Salles T, et al. A genetic

programming approach for feature selection in highly dimensional skewed data. Neurocomputing. 2018;**273**: 554-569

[19] Guyon I, Elisseeff A. An introduction to variable and feature selection. Journal of Machine Learning Research. 2003;**3**(March):1157-1182

[20] Shri TKP, Sriraam N. Spectral entropy feature subset selection using sepcor to detect alcoholic impact on gamma sub band visual event related potentials of multichannel electroencephalograms (EEG). Applied Soft Computing. 2016;**46**:441-451

[21] Almaghrabi F. Machine learning methods for predicting traumatic injuries outcomes [PhD thesis]. The University of Manchester; 2020

[22] Almaghrabi F, Xu DL, Yang JB. Features selection and improving for trauma outcomes prediction models. In: Data Science and Knowledge Engineering for Sensing Decision Support. Singapore: World Scientific Publishing Co. Pte. Ltd.; 2018. pp. 1309-1314

[23] Hira ZM, Gillies DF. A review of feature selection and feature extraction methods applied on microarray data. Advances in Bioinformatics. 2015;**2015**

[24] Hua J, Tembe WD, Dougherty ER. Performance of feature-selection methods in the classification of high-dimension data. Pattern Recognition. 2009;**42**(3):409-424

[25] Ferreira AJ, Figueiredo MRAT. Efficient feature selection filters for high-dimensional data. Pattern Recognition Letters. 2012;**33**(13): 1794-1804

[26] Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007; **23**(19):2507-2517

[27] Mumtaz W, Vuong PL, Xia L, Malik AS, Rashid RBA. An EEG-based machine learning method to screen alcohol use disorder. Cognitive Neurodynamics. 2017;**11**(2):161-171

[28] Mumtaz W, Kamel N, Ali SSA, Malik AS, et al. An EEG-based functional connectivity measure for automatic detection of alcohol use disorder. Artificial Intelligence in Medicine. 2018;**84**:79-89

[29] Neagoe V-E, Carata S-V. Subject independent drunkenness detection using pulse-coupled neural network segmentation of thermal infrared facial imagery. In: Proceedings of the 5th International Conference on Applied and Computational Mathematics. Sofia: IARAS; 2016. pp. 305-312

[30] Wan L, Baldridge RM, Colby AM, Stanford MS. Association of p3 amplitude to treatment completion in substance dependent individuals. Psychiatry Research. 2010;**177**(1–2): 223-227

[31] Kranenburg RF, Peroni D, Affourtit S, Westerhuis JA, Smilde AK, van Asten AC. Revealing hidden information in GC–MS spectra from isomeric drugs: Chemometrics based identification from 15 eV and 70 eV EI mass spectra. Forensic Chemistry. 2020; **18**:100225

[32] Mahmud MS, Fang H, Wang H, Carreiro S, Boyer E. Automatic detection of opioid intake using wearable biosensor. In: 2018 International Conference on Computing, Networking and Communications. Maui, USA: IEEE; 2018. pp. 784-788

[33] Song F, Mei D, Li H. Feature selection based on linear discriminant analysis. In: 2010 International Conference on Intelligent System Design and Engineering Application. Vol. 1. Changsha, China: IEEE; 2010. pp. 746-749

[34] Feldesman MR. Classification trees as an alternative to linear discriminant analysis. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists. 2002;**119**(3):257-275

[35] Langan RT, Archibald RK, Lamberti VE. Nuclear forensics analysis with missing data. Journal of Radioanalytical and Nuclear Chemistry. 2016;**308**(2):687-692

[36] Brown RL. Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of five methods. Structural Equation Modeling: A Multidisciplinary Journal. 1994;**1**(4):287-316

[37] Graham JW, Hofer SM, MacKinnon DP. Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. Multivariate Behavioral Research. 1996;**31**(2):197-218

[38] Rubin DB. Inference and missing data. Biometrika. 1976;**63**(3):581-592

[39] Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: What is it and how does it work? International Journal of Methods in Psychiatric Research. 2011;**20**(1): 40-49

[40] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B: Methodological. 1977;**39**(1):1-22

[41] Stekhoven DJ, Bühlmann P. Missforest—Non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;**28**(1):112-118

[42] Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using mice: A CALIBER study. American Journal of Epidemiology. 2014;**179**(6):764-774

[43] Ding Y, Ross A. A comparison of imputation methods for handling missing scores in biometric fusion. Pattern Recognition. 2012;**45**(3):919-933

[44] Deng Y, Chang C, Ido MS, Long Q. Multiple imputation for general missing data patterns in the presence of high-dimensional data. Scientific Reports. 2016;**6**(1):1-10

[45] Ross Quinlan J. C4. 5: Programs for Machine Learning. San Mateo, California: Elsevier; 2014

[46] Azar AT, El-Metwally SM. Decision tree classifiers for automated medical diagnosis. Neural Computing and Applications. 2013;**23**(7–8):2387-2403

[47] Vuong TP, Loukas G, Gan D, Bezemskij A. Decision tree-based detection of denial of service and command injection attacks on robotic vehicles. In: 2015 IEEE International Workshop on Information Forensics and Security. Rome, Italy: IEEE; 2015. pp. 1-6

[48] Lolli F, Ishizaka A, Gamberini R, Balugani E, Rimini B. Decision trees for supervised multi-criteria inventory classification. Procedia Manufacturing. 2017;**11**:1871-1881

[49] Greco S, Matarazzo B, Słowiński R. Decision rule approach. In: Multiple Criteria Decision Analysis. New York: Springer; 2016. pp. 497-552

[50] Letham B, Rudin C, McCormick TH, Madigan D, et al. Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. The Annals of Applied Statistics. 2015;**9**(3):1350-1371

[51] Karabiyik U, Aggarwal S. Advanced automated disk investigation toolkit. In:

IFIP International Conference on Digital Forensics. Cham: Springer; 2016. pp. 379-396

[52] Xu D-L, Liu J, Yang J-B, Liu G-P, Wang J, Jenkinson I, et al. Inference and learning methodology of belief-rule-based expert system for pipeline leak detection. Expert Systems with Applications. 2007;**32**(1):103-113

[53] Sachan S, Yang J-B, Xu D-L, Benavides DE, Li Y. An explainable AI decision-support-system to automate loan underwriting. Expert Systems with Applications. 2020;**144**:113100

[54] Yang J-B, Xu D-L. Evidential reasoning rule for evidence combination. Artificial Intelligence. 2013;**205**:1-29

[55] Labreuche C, Fossier S. Explaining multi-criteria decision aiding models with an extended Shapley value. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. California: AAAI Press; 2018. pp. 331-339

[56] Zhong Q, Fan X, Luo X, Toni F. An explainable multi-attribute decision model based on argumentation. Expert Systems with Applications. 2019;**117**: 42-61

[57] Pessach D, Singer G, Avrahami D, Ben-Gal HC, Shmueli E, Ben-Gal I. Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. Decision Support Systems. 2020:113290

[58] Zhan H, Cao Y. Relationship explainable multi-objective reinforcement learning with semantic explainability generation. arXiv preprint arXiv:1909.12268. 2019

[59] Antonelli M, Bernardo D, Hagras H, Marcelloni F. Multiobjective evolutionary optimization of type-2 fuzzy rule-based systems for financial data classification. IEEE Transactions on Fuzzy Systems. 2016;**25**(2):249-264

[60] Evans BP, Xue B, Zhang M. What's inside the black-box? A genetic programming method for interpreting complex machine learning models. In: Proceedings of the Genetic and Evolutionary Computation Conference. New York: Association for Computing Machinery; 2019. pp. 1012-1020

[61] Piltaver R, Luštrek M, Zupančič J, Džeroski S, Gams M. Multi-objective learning of hybrid classifiers. In: Proceedings of the Twenty-First European Conference on Artificial Intelligence. Amsterdam: IOS Press; 2014. pp. 717-722

[62] Jiang L, Liu S, Chen C. Recent research advances on interactive machine learning. Journal of Visualization. 2019;**22**(2):401-417

[63] Spinner T, Schlegel U, Schäfer H, El-Assady M. ExplAIner: A visual analytics framework for interactive and explainable machine learning. IEEE Transactions on Visualization and Computer Graphics. 2019;**26**(1): 1064-1074

[64] Liu S, Bremer PT, Thiagarajan JJ, Srikumar V, Wang B, Livnat Y, et al. Visual exploration of semantic relationships in neural word embeddings. IEEE Transactions on Visualization and Computer Graphics. 2017;**24**(1):553-562

[65] Chen Y, Xu P, Ren L. Sequence synopsis: Optimize visual summary of temporal event data. IEEE Transactions on Visualization and Computer Graphics. 2017;**24**(1):45-55

[66] El-Assady M, Sevastjanova R, Sperrle F, Keim D, Collins C. Progressive learning of topic modeling parameters: A visual analytics framework. IEEE Transactions on

Visualization and Computer Graphics. 2017;**24**(1):382-391

[67] Sacha D, Zhang L, Sedlmair M, Lee JA, Peltonen J, Weiskopf D, et al. Visual interaction with dimensionality reduction: A structured literature analysis. IEEE Transactions on Visualization and Computer Graphics. 2016;**23**(1):241-250

[68] Shrikumar A, Greenside P, Kundaje A. Learning important features through propagating activation differences. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR.org. United States: PMLR; 2017. pp. 3145-3153

[69] Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS One. 2015; **10**(7)

[70] Berk Ustun, Stefano Traca, Cynthia Rudin. Supersparse linear integer models for interpretable classification. arXiv preprint arXiv:1306.6677. 2013

[71] Xu K, Ba J, Kiros R, Cho K, Courville A, Salakhudinov R, et al. Show, attend and tell: Neural image caption generation with visual attention. In: International Conference on Machine Learning. United States: PMLR; 2015. pp. 2048-2057

[72] Friedman JH. Greedy function approximation: A gradient boosting machine. Annals of Statistics. 2001: 1189-1232

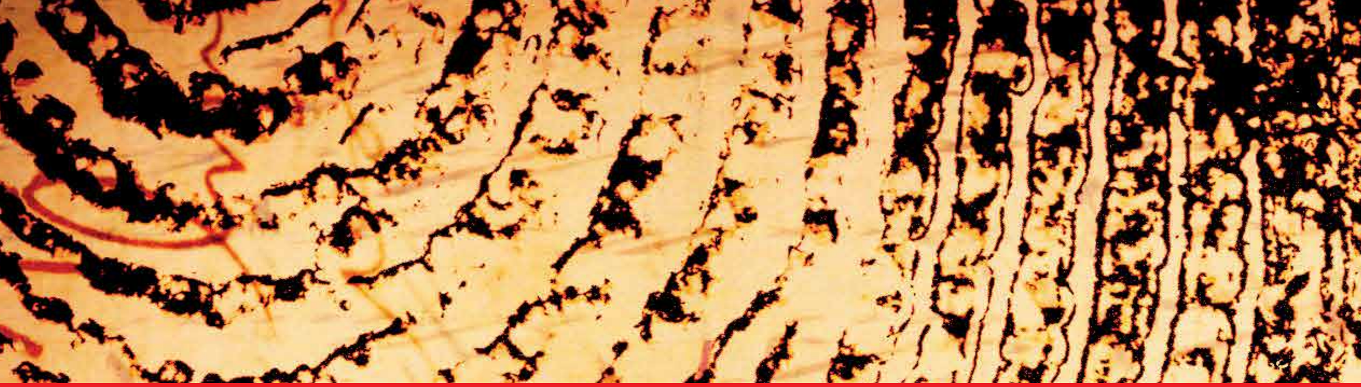
[73] Daniel W Apley, Jingyu Zhu. Visualizing the effects of predictor variables in black box supervised learning models. arXiv preprint arXiv: 1612.08468. 2016

[74] Goldstein A, Kapelner A, Bleich J, Pitkin E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. Journal of Computational and Graphical Statistics. 2015;**24**(1):44-65

[75] Fisher A, Rudin C, Dominici F. Model class reliance: Variable importance measures for any machine learning model class, from the "rashomon" perspective. 2018;**68**. arXiv preprint arXiv:1801.01489

[76] Al-Nemrat A, Benzaid C. Cybercrime profiling: Decision-tree induction, examining perceptions of internet risk and cybercrime victimisation. In: 2015 IEEE Trustcom/BigDataSE/ISPA, Volume 1. Helsinki, Finland: IEEE; 2015. pp. 1380-1385

[77] Yang M, Chow K-P. An information extraction framework for digital forensic investigations. In: IFIP International Conference on Digital Forensics. Orlando, FL,USA: Springer; 2015. pp. 61-76



Edited by B Suresh Kumar Shetty
and Pavanchand Shetty H

It is our pleasure to place before you the book *Digital Forensic Science*. This book makes up a major part of the broad specialty of Digital Forensic Science, comprising mainly of tools and technologies of cyber forensic experts for their future practice. This book has been designed to merge a range of new ideas and unique works of authors from topics like fundamental principles of forensic cyber analysis, and protocols and rules needed for the best digital forensics. We hope that it will be useful to practitioners of forensic medicine, experts, cyber experts, law makers, investigating authorities, and undergraduate and postgraduate medical school graduates of medicine.

Published in London, UK
© 2020 IntechOpen
© JanakaMaharageDharmasena / iStock

IntechOpen

ISBN 978-1-83968-742-6
9 781839 687426