1. Jamming Attacks Detection and Classification in Cognitive UAV Radios

Ali Krayani, Lucio Marcenaro, Carlo Regazzoni

Università degli studi di Genova
DITEN, The Electrical, Electronics and Telecommunication Engineering and Naval Architecture Department

Italian National Inter-University Consortium for Telecommunications (CNIT)

Abstract: *Unmanned Aerial Vehicles (UAVs) attracted both industry and research community owing to their fascinating features like mobility, deployment flexibility and strong Line of Sight (LoS) links. The integration of Cognitive Radio (CR) can greatly help UAVs to overcome several issues especially spectrum scarcity. However, the dynamic radio environment in CR and the strong dependence of safe communications from LoS channels integrity in UAV communications make the Cognitive-UAV-Radio vulnerable to jamming attacks. This chapter presents a novel method for joint detection and automatic classification of multiple jammers attacking with different modulation schemes. The method is based on learning a representation of the radio environment encoded in a Generalized Dynamic Bayesian Network (GDBN), whilst multiple GDBN models represent various jamming signals under different modulation schemes. The simulated results demonstrate that the presented GDBN-based method outperforms Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) and Stacked Autoencoder (SAE) in terms of classification accuracy and achieves a higher degree of explainability of its own decisions by interpreting causes and effects at hierarchical levels under the Bayesian learning and reasoning processes.*

1.1 Introduction

Unmanned Aerial Vehicles (UAVs) have attracted the attention of the telecommunication community and industry due to their remarkable features such as deployment flexibility, mobility, and dominant Line of Sight (LoS) links [3]. UAVs are expected to play a significant role in future wireless communications (beyond 5G) and are regarded as an important complement to the terrestrial networks from the sky [4]. However, UAV-based communications will face several problems, such as spectrum scarcity due to the explosively increasing number of connected UAVs in various applications like smart cities, surveillance and IoT [5], energy efficiency due to the on-board limited battery lifetime [6, 7], and physical layer security (e.g., jamming attacks) due to the open nature of ground-to-air wireless channels and the dominant LoS propagation links [8].

Cognitive Radio (CR) can provide a promising solution for UAVs that can tackle the aforementioned problems and achieve the capability of reaching and maintaining connectivity with a

This report contains content that has been published in [1, 2].

2. Physical and MAC Layer Techniques for Secure Positioning in Navigation Satellite Systems

Francesco Ardizzon, Laura Crosara, Nicola Laurenti, Stefano Tomasin

Università degli Studi di Padova

Abstract: Global navigation satellite systems (GNSSs) have become crucial for many applications, both military and civil. Being a widespread technology, there are several motivations to mount attacks against it. To counter these threats, in this chapter, we present GNSS signal authentication techniques operating at both physical and data layers. After a brief introduction of GNSS based positioning, we introduce the main attacks and threats against such a service. Next, we present possible countermeasures, distinguishing between authentication at the system and receiver side.

Next, we discuss the use of cross-checks between authenticated and open GNSS observables: these techniques aim at improving the navigation performance by enlarging the set of signals used to compute the position, velocity, and time (PVT), using as anchors the signals authenticated by services such as chips-message robust authentication (CHIMERA) or commercial authentication service (CAS). Finally, we show how a spoofer can tamper the timing information retrieved from the PVT solution computed by a receiver, in a time-spoofing attack. Lastly, we discuss timing assurance techniques.

2.1 Introduction

Since the 1980s, when the global positioning system (GPS) became operational as the first global navigation satellite system (GNSS), its applications and use have rapidly increased. Nowadays, almost every one of us typically has one or more devices to move around with the support of GNSSs. Our lives are becoming increasingly reliant on precise positioning and timing, as technological advances make GNSS devices more affordable. Indeed, precise positioning is critical in industry sectors such as surveying, construction, and logistics for automation and safety applications.

Signals broadcast by satellites include two main components, the *navigation data* and the *ranging code*. The first is a binary signal containing information necessary to compute a position, velocity, and time (PVT) solution at the receiver, such as satellite ephemeris (position and velocity), clock bias parameters, almanac, satellite health status, and other supplementary data. The latter consists of a binary code, called pseudo-random noise (PRN) sequence, different for each satellite and orthogonal to each other, used to identify signals of different satellites transmitted simultaneously in a code division multiple access (CDMA) fashion. The PRN code is used at the receiver, when computing the PVT, for estimating its distances to the satellite. The navigation data is modulo-2 added to the binary high-rate and periodic PRN sequence, obtaining a spread-spectrum signal that modulates the system carrier. For example, the data rate of Galileo's open system (OS) dissemination channel is 120 bit/s; that of GPS C/A and L1C is 50 bit/s, as for the

3. Physical Layer Security and Authentication with Practical Transmissions

Linda Senigagliesi, Marco Baldi, Ennio Gambi, Franco Chiaraluce

Università Politecnica delle Marche
Via Brecce Bianche 12, Ancona, Italy

Abstract: *Classic approaches to communications security neglect the physical layer and rely on computational security solutions working at higher layers, like cryptography. However, the physical layer of communications provides for some interesting properties that may be significant for security, like uniqueness, randomness and reciprocity. In this chapter we give an overview of techniques that allow achieving confidentiality and authentication at the physical layer taking into account the typical constraints of practical transmissions, like finite-length coding and discrete modulation formats. We show that, even with these constraints in mind, the physical layer can be a source of security for communications, both in terms of confidentiality and authentication, and provides a basis for hardening security at higher layers.*

3.1 Introduction

The whole area of telecommunications is facing an increasing need for security, intended both as confidentiality of exchanged data and authentication of users and messages. Two classical approaches exist to communication security, ideally represented, respectively, by the two pioneers Alan Touring and Claude Elwood Shannon: *computational security* and *unconditional security*, with their own advantages and disadvantages.

The former paradigm is at the basis of cryptography, which aims at protecting legitimate users' data from a computationally constrained attacker [1], neglecting the effects of transmission. This implies that the overall security of the system depends on the computational resources available to attackers. Opposedly, unconditional security, commonly known as physical layer security (PLS), relies on the intrinsic differences of communications channels to differentiate legitimate users from attackers, without putting any constraints on the computing power of attackers [2]. Originally introduced for achieving confidentiality between two legitimate users (Alice and Bob) in the presence of a passive eavesdropper (Eve), or wiretapper [3], the principles of unconditional security can also be exploited to achieve user authentication, that is, hysical layer authentication). For an overview of PLS and physical layer authentication techniques the interested reader is referred to [4, 5].

The main advantage of PLS and physical layer authentication is related to the fact that the security of confidentiality and authentication does not rely on any pre-shared secret, and is also independent of the eavesdropper's computational power. Still, PLS and physical layer authentication are not sufficient in several practical contexts, due to the impossibility of taking into account all practical aspects of communication channels and techniques, as well as because of the chance for the attackers to implement active attacks that can compromise security. For this

4. Sensing and Privacy in Future Telecommunication Networks: a Joint Perspective

Francesco Gringoli, Renato Lo Cigno, Marco Cominelli, Lorenzo Ghiro

University of Brescia and CNIT (Consorzio Nazionale Interuniversitario Telecomunicazioni), Italy

Abstract: *Wireless communications leverage the description of the propagation channel through the CSI to achieve unprecedented transmission speed and efficiency. orthogonal frequency division multiplexing modulations simplify the collection of the CSI and massive multiple-input multiple-output technologies empowers spatial reuse of the spectrum and multi-stream transmissions. All of this is rooted in the short-term stability of the propagation channel, which allow almost perfect equalization based on the CSI. The long term behavior of the channel, well beyond the coherence time, is normally not useful for transmission performance, but yields useful information on the environment leading, in the past decade, to the idea of* ***Joint Communication and Sensing (JCS)****. Sensing includes people localization and tracking, object recognition, but also e-health and emotional computing applications. We stress that this is very different from standard device localization as the the ambient embeds its own characteristics in the communications signals, which can be used as a sort of radar, or more appropriately a scanner that carries some information on all the surfaces and obstacles that it has encountered in its multipath travel between the transmitter an one or more receivers. The potential of JCS is enormous, and all future telecommunication systems, from 5G and 6G to emerging Wi-Fi standards foresee JCS in their evolution path. At the same time this potential entails also enormous privacy and security/safety concerns, as JCS is a pervasive tracking system that cannot be countered with any cryptographic means, as the information is embedded within the signal at the physical level, and not on the data carried by the modulated signal. This chapter collects and reorders research and results we have obtained in the past few years trying to understand what are the key characteristics that actually empower JCS, but most of all to study techniques that can prevent unauthorized sensing, thus protecting users' privacy and safety. Sensing results are mostly based on Angle Information (AI) and aximum likelihood) techniques, and proper conceptual models explaining how and why it works, and allowing predictions on its reliability and dependability are still not available. Yet, some of these AI/ML techniques are robust enough to be reproducible in different scenarios and yield stable and consistent results, thus it is very important to understand how to counter them. We call this countering activity* ***obfuscation****, to highlight its basic property of hiding sensing information embedded in the Expectation-maximization (EM) signal, while preserving communication capability and performance. Results we report in this work show that obfuscation is feasible and it can be included in protocols that allow legitimate sensing while preventing attacks and unauthorized used of it. Furthermore, we discuss future directions and foundational research we deem needed to fully exploit JCS potential and protect users.*

5. Location Security in the Digital Era: Threats Detection, Analysis, and Mitigation Strategies

Stefania Bartoletti[1], Ivan Palamà[1], Giuseppe Bianchi[1], Nicola Blefari Melazzi[1]

Danilo Orlando[2]

[1]Università di Roma, Tor Vergata, Italy
[2]Università Telematica Unicusano, Italy

Abstract: *In the incoming fifth generation mobile communication network (5G), localization services, which in the past were mainly provided by non-cellular technologies, are now combined with solutions based upon cellular technologies and integrated within 5G architecture [8]. As a consequence, 5G architecture can improve the performance of the localization system but at the cost of new security vulnerabilities that need to be investigated and mitigated. This chapter develops functions and algorithms to detect attacks over the air interface, providing alerts that can be used for mitigation purposes.*

5.1 Introduction

The security of location data in 5G networks is a critical aspect as location data is used for a wide range of applications such as emergency services, navigation, and location-based services. The 5G networks use a combination of 3GPP technologies such as the New Radio (NR) uplink and downlink reference signals, as well as non-3GPP technologies such as Global Navigation Satellite System (GNSS), Terrestrial Beacon Systems (TBS), WLAN-based localization services, and other sensors to extract location metrics. However, these technologies also bring new security challenges as the location data is highly vulnerable to spoofing and meaconing attacks caused by malicious actors.

Thus, it becomes of primary importance the design of a generic location security function with the aim of detecting and suitably handling any deviation from the true locations to mitigate these threats. This function should consist of two stages: a preliminary stage for the detection of an attack and a second stage aimed at handling the fake measurements generated by the malicious actor. A high-level description of the designed location security function is shown in Figure 5.1.

Additionally, the location security function must be able to manage and perform data fusion of measurements provided by heterogeneous sensors. As shown in Figure 5.1, this function must be able to integrate and process data from multiple sensors to ensure the accuracy and integrity of the location data.

In conclusion, securing location data in 5G networks is a complex task that requires a multi-faceted approach. It involves the integration of advanced security features, the implementation of a generic location security function, and the ability to manage and process data from multiple

6. A Cybersecurity Framework for Securing Digital Service Chains

Giovanni Grieco, Domenico Striccoli, Matteo Repetto, Alessandro Carrega, Luigi Alfredo Grieco, Raffaele Bolla, Giuseppe Piro, Gennaro Boggia

DEI, Politecnico di Bari. Via Orabona 4, Bari, Italy

Institute for Applied Mathematics and Information Technologies (IMATI), CNR, Genoa

CNIT, Consorzio Nazionale Interuniversitario per le Telecomunicazioni

DITEN, University of Genoa, Genoa, Italy

S2N National Lab, CNIT, Genoa, Italy

Abstract: *Today, the digital economy is pushing new digital services and digital service chains through the interconnection of processes and services across different domains and organizations. In such a scenario, an architecture is needed that effectively fulfills all the main security issues: mutual trustworthiness of entities in partially unknown topologies, identification and mitigation of advanced multi-vector threats, management and propagation of sensitive data, and advanced identity management and access control procedures. Based on these considerations, this chapter aims to reach two goals. First, it proposes a new methodological approach by designing a framework that implements heterogeneous security services for distributed systems that combine together digital resources and components from multiple domains. The framework focuses on three novel aspects: i) full automation of the processes that manage the whole system, ii) dynamic adaptation of operations and security tasks to newest attack patterns, and iii) real-time adjustment of the level of detail of inspection and monitoring processes. Second, it proposes an authentication and authorization module that automatically protects the information flowing among the framework modules, guaranteeing resource availability only to authenticated subjects. Experimental tests show that the proposed module enables authentication and authorization procedures, while maximizing the flexibility of the set of access control policies and providing an efficient service protection.*

6.1 Introduction

The most remunerative business in the digital economy will be the creation of value chains for processing data, through the interconnection of processes, products, services, software, and things from multiple vendors on a growing scale. Fully-automated software and environments will evolve and morph during run-time, without the explicit control of software engineers [1].

The uptake of cloud services and IoT has raised the interest in combining together digital resources and components from multiple domains and locations, to create Cyber-Physical Systems

7. Internet of Things Security Issues in LoRaWAN and Bluetooth Low Energy

Andrea Lacava, Pierluigi Locatelli, Pietro Spadaccino, Francesca Cuomo

Research Unit CNIT-University of Rome Sapienza

Abstract: *The Internet of things (IoT) technologies are attracting the interest of scientific and industrial communities given the huge number of applications that can be designed with energy-limited devices connected to the Internet. LoRaWAN (Long Range Wide Area Network), a data-link layer with long range, low power, and low bit rate, appeared as a promising solution for IoT in which, end-devices communicate with gateways through a single hop at long distances. On the other side, the short-range IoT can leverage the potentialities of Bluetooth Low Energy (BLE) in its mesh network setting to interconnect, in proximity, multiple devices. Both technologies are already hitting a large market, but there are still several issues dealing with their security. This chapter is dedicated to the analysis of selected security issues affecting IoT networks and specifically low-power wireless systems such as LoRaWAN and BLE.*

7.1 Introduction

The approaching IoT era will lead billions of entities to be simultaneously connected, opening new challenges regarding network management and data exchange rules. While the goal of IoT to realize an environment within which things are uniquely identified and able to interact with one another through the exchange of information seems really close, the concerns about security and privacy still remain far to be solved [1].

The exponential increase in the number of agents in the network determines an equivalent increase in vulnerabilities that can be exploited by criminals to gain access and control of the systems. In addition, the use of technologies that have not yet followed an official standardization process and therefore do not have full support from open-source communities makes it even more urgent to raise awareness, introduce cybersecurity concepts and ideas and finally create defense systems specifically designed to act in conditions where energy consumption is a concern and where there is no use of IP stacks, such as IoT mesh intranets.

This work is dedicated to the analysis of selected security issues affecting the IoT networks and specifically low-power wireless systems such as LoRaWAN and BLE. These technologies represent the long-range and the short-range most used wireless networks, respectively, thus giving a full overview of the security criticalities in the spectra. LoRaWAN is a MAC protocol targeted for battery-operated devices working on top of the proprietary Long Range (LoRa) radio modulation and represents one of the most promising Low Power Wide Area Network (lpwan) technologies, with its coverage significantly expanding in recent years.

BLE is a wireless, low-power personal area network that operates in the 2.4 GHz ISM band. Starting from single pairing applications such as connections between smartphones and IoT devices, the BLE protocol is gaining new momentum thanks to its Mesh Service Models [2] that

8. Trustworthy Task Allocation in IoT: a Cognitive Game-Theoretical Use Case

Virginia Pilloni, Marco Martalò, Luigi Atzori

Networks for Humans (Net4U) Lab
Department of Electrical and Electronic Engineering, University of Cagliari
National Telecommunication Inter University Consortium, Research Unit of Cagliari
09123 Cagliari, Italy

Abstract: *This Chapter analyzes how malicious attacks affect the performance of a heterogeneous Internet of Things (IoT) system where cognitive devices collaborate to negotiate task assignments. In the reference scenario, the involved devices create clusters, each managed by a Cluster Head (CH). Whenever a task is required, the CH triggers spectrum sensing to detect spectrum holes that can be opportunistically exploited by the nodes of the cluster for task allocation. In this scenario, not all the nodes are Honest Nodes (HN). Indeed, Malicious Nodes (MNs) may hinder the process and try to disrupt it by providing tampered data, which would lead to a higher likelihood that the spectrum sensing is not performed correctly. When the spectrum is considered free, the cluster nodes negotiate to execute the required task by means of an auction-based game theory approach. The negotiation takes into account two factors: the reward gained from contributing to the execution of the task, which is provided to the node that wins the competition, and the energy cost to perform the task. Specifically, the Chapter investigates how MNs affect the reward aspect when they try to gain maximum control over the task and potentially launch a Denial of Service (DoS) attack. Extensive simulations are run to assess the effect of the key system parameters on the overall performance and provide recommendations for future research.*

8.1 Introduction

In Wireless Mesh Networks (WMNs) nodes can communicate with one another wirelessly, forming a dynamic and non-hierarchical network topology. WMNs can also connect to the external Internet through designated gateways, as documented in previous research [1]. This communication technology can be useful in Internet of Things (IoT) contexts, where resource-constrained devices, such as sensors, personal electronics, and smart vehicles [2], cooperate to accomplish specific application tasks, such as transmitting data collected from the physical world, computing relevant data quantities, storing data, etc. [2]. Indeed, many IoT applications involve the widespread collection and aggregation of data from various devices in order to monitor and manage the physical world using short-range communications.

Similar to WMNs, the IoT can rely on clusters managed by Cluster Heads (CHs) to establish a hierarchy among the participants in the communication infrastructure. CHs are devices that have

9. Video Deepfake Detection by Identity Analysis

Davide Cozzolino, Giovanni Poggi, Luisa Verdoliva

Università degli Studi di Napoli Federico II
Consorzio Nazionale Interuniversitario per le Telecomunicazioni

Abstract: *Face manipulation technology is advancing very fast, and new synthetic generation methods are being proposed day by day. They can be used to create hyper-realistic manipulated videos, also known as deepfakes, in which the face of one individual is replaced with another one (face swapping) or the expressions and head movements are controlled by another person's (facial reenactment). While there are many positive applications of these tools in diverse applications, such as movie production and the video game industry, they can also be used maliciously to manipulate public opinion during elections, discredit or blackmail people, and even create non-consensual pornography and disinformation campaigns. In the last few years, a large number methods have been proposed to detect deepfakes. However, they are mostly supervised and need be trained on the same manipulation methods present in the test videos. Therefore, they exhibit poor generalization ability across different types of facial manipulations, e.g., from face swapping to facial reenactment. This chapter presents part of the research activities on video deepfake detection developed at the University of Napoli Federico II (in collaboration with the Technical University of Munich) and describes a detection strategy that can cope with the wide variety of manipulation methods and scenarios encountered in the real world. The main idea is to expose video manipulations by detecting anomalies in the soft biometric traits of the portrayed individual, such as temporal facial features or how a person moves and talks, which are peculiar of that specific identity. As a consequence, the detector is trained only on real videos, with no need of fake videos of any kind. This is a major strength as it ensures a natural generalization to new and unseen types of attack. Moreover, the use of high-level semantic features guarantees robustness to widespread and disruptive forms of post-processing, such as compression or resizing. The experimental analysis confirms the generalization and robustness property of this approach, with significant improvements with respect to the current state-of-the-art. Code and trained network of our work are publicly available*[1].

9.1 Introduction

With the advancement of synthetic media generation technology, there is a steady increase in the level of photorealism, as more and more video manipulation methods emerge. This phenomenon raises serious concerns about the trustworthiness of media content and the diffusion of fake news over the web. Especially, deep learning-based methods for image, video and audio synthesis and manipulation, so called deepfakes, are widely spread and more and more effective. With reference to videos, faces can be altered in various ways that can be broadly classified in two categories: (a) face swapping, where a face in a video is replaced with another person's face [1, 2], (b) facial reenactment, in which a target video is modified using a source video from which it inherits the

[1] https://github.com/grip-unina/id-reveal

10. Detecting Audio Deepfakes using Semantic Traces

Davide Salvi, Clara Borrelli, Sara Mandelli, Paolo Bestagini, Stefano Tubaro

Image and Sound Processing Lab (ISPL)
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy

Abstract: *Recent advances in deep learning and computer vision have made the synthesis and counterfeiting of multimedia content more accessible than ever, leading to possible threats and dangers from malicious users. New media categories have arisen, such as deepfakes: synthetic multimedia content generated through deep learning techniques capable of synthesizing a target person's identity or biometric aspects. The malicious use of deepfakes can pose multiple risks in several domains, including the audio one, which is experiencing growth in the development of synthetic speech generation methods. This has prompted the development of synthetic speech detection algorithms to counteract potential malicious uses, and prevent cases of fraud or identity theft. However, the high quality of generated speech data has shown that traditional low-level features are no longer adequate for detecting such data, requiring the use of more advanced techniques. In this chapter, we explore the use of high-level features that extract semantic traces from data to perform speech deepfake detection. In particular, we present a detector that exploits the emotional content of speech and another one based on the joint use of prosody features and speaker identity cues. The rationale behind the proposed methods is that audio deepfake techniques cannot correctly synthesize natural human behavior. The considered approaches proved to be highly robust in the speech deepfake detection task, confirming the validity of our initial hypothesis.*

10.1 Introduction

Thanks to the constant development of new technologies and the massive evolution of neural networks, synthetic speech generation is nowadays an effortless operation and it is becoming increasingly difficult to distinguish synthetic audio material from original one. While this opens the door to new challenging and stimulating scenarios, it can also lead to problematic situations.

For instance, deepfakes have been recently used with malicious intent in various cases, especially in mass and social media. Some examples concern the spreading of fake news [1] and fraud cases [2], which led to ethical considerations regarding the use of artificial intelligence [3].

Moreover, impersonation attacks may be dangerous in everyday life scenarios. Voice signal is often used to assess the identity or control devices through speech-human interfaces. The availability of synthesis techniques able to reproduce any voice puts at risk the reliability of such systems.

Synthetic speech may be problematic also in scenarios involving forensic investigations. Among other pieces of evidence like photographs or video frames, the voice recordings, their transcriptions and the recording context may be crucial. It is easy to imagine a scenario where an audio evidence is maliciously forged to simulate, for example, a conversation that never happened.

11. A Hybrid Architecture for the Classification and Localization of GAN-generated Images with Improved Robustness and Generalization Capabilities

Jun Wang, Benedetta Tondi, Mauro Barni

University of Siena
Department of Information Engineering and Mathematics

Abstract: *Generative Adversarial Network (GAN) models are nowadays able to generate synthetic images which are visually indistinguishable from the real ones, thus raising serious concerns about the spread of fake news and the need to develop tools to distinguish fake and real images in order to preserve the trustworthiness of digital images. In this report, we present a hybrid deep learning architecture for the detection/classification that includes a localization branch, devoted to the identification of the image regions manipulated by GANs. Even if our goal is the detection/classification we found that adding a localization branch helps the network to focus on the most relevant image regions, with significant improvements in terms of generalization capabilities and robustness against image processing operations. The effectiveness of the proposed hybrid network was validated on two new image forensic tasks, namely, the detection of fake images of climate change (flood images) generated by the ClimateGAN architecture and the classification of GAN face editing.*

11.1 Introduction

Multimedia Forensics (MF) has received a constantly increasing attention over the last two decades. More recently, the remarkable achievements in image synthesis obtained by Generative Adversarial Networks (GANs) [1] has added a further challenge for the MF community. Nowadays, GAN architectures can synthesise artificial images with a quality that can easily fool non-professional users. As a consequence, MF researchers have increased their efforts to develop tools capable of detecting GAN-generated images and videos, with particular emphasis on the detection of synthetic face images [2, 3, 4] and videos [5, 6, 7], usually referred to as deepfakes.
Traditionally, the pipeline for the design of an effective image forgery detector consists in extracting a set of hand-crafted features, such as color differences [29], saturation cues [30], or corneal highlights [31], and use them to detect the presence of forgery artefacts. By focusing on the detection of GAN-generated images and forgeries, some works have proposed to analyze the frequency artefacts introduced by GANs due to the upsampling operation applied by such networks [32, 33]. In parallel to methods based on handcrafted features, data driven detectors based on various CNN architectures, like XceptionNet [5], ResNet [34], and EfficientNets [42] are gaining more and more popularity, due to their unsurpassed performance, at least when the test conditions match those used during training.
Forgery localization has received relatively less attention, yet the possibility to localize where an

12. Container and Content-based Media Signatures for Open-World Multimedia Forensics

Daniele Baracchi[1,2], **Dasara Shullani**[1,2], **Andrea Montibeller**[1,3], **Massimo Iuliani**[1,2]
Cecilia Pasquini[1,4,5], **Giulia Boato**[1,3], **Alessandro Piva**[1,2], **Francesco De Natale**[1,3]

[1]CNIT - National Inter-University Consortium for Telecommunications, Parma
[2]Department of Information Engineering, University of Florence
[3]Department of Information Engineering and Computer Science, University of Trento
[4]Fondazione Bruno Kessler, Trento
[5]Futuro e Conoscenza srl, Roma

Abstract: *The detection of image and video deceptions is getting more and more relevant in several fields such as investigation, intelligence and forensics. Multimedia forensics researchers keep designing new tools and updating available detectors to understand the processing the media has been subjected to. Even though these tools can be effectively used under controlled environments, they are generally unreliable in open-world settings, where the investigated content may have undergone several unknown processing. In this chapter, we present a novel framework to discriminate different toolchains of media manipulation and processing. We introduce the concept of media signature encoding to map visual contents in spaces where media corresponding to similar processing toolchains cluster together. We also demonstrate that this property still holds for toolchains that are not known when building the encoder, thus making the framework applicable in open-world settings where the forensic analyst can face both known and unknown manipulations. The proposed framework has been tested in a challenging experimental setup involving manipulated videos, and the results show that encoded signatures are effective in determining whether a video sequence under analysis belongs to a known life cycle or to a never seen processing toolchain, and if a subset of media items share the same history. This framework can be considered a first step towards the use of forensic features to characterize media life cycles in open-world settings.*

12.1 Introduction

Massive amounts of visual data are uploaded every day to social media platforms by nearly 4 billion active users. According to recent estimates, 2 million hours of video are uploaded to YouTube every minute[1]. The reason behind the popularity of sharing images and videos is actually rooted in the structure of the human brain, which is extremely fast and efficient at processing visual information (as opposed to, e.g., textual content). The result is that visual media grab more attention from users, engage them more and increase significantly the propensity to share. Visual data are responsible for the viral diffusion of information through social media and web channels, and they play a key role in the digital life of individuals and societies.

[1]https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/