Digital identity and identification: Secure remote identity verification
- CODE 2021 Workshop Report-

Organizer and Moderator:

Dr. Andreas Wolf (Bundesdruckerei, Andreas.Wolf@BDR.de)

Abstract

The secure remote verification of identities is becoming increasingly important. This process has already been observed over the past few years. The COVID-19 pandemic was one reason for the growing relevance of remote authentications that developed much faster than originally expected. Once authentication credentials have been issued, they can be used for further authentication. But where can such credentials be derived from? In many cases, this is done on the basis of video sessions, in which people and their identification documents can be inspected by remote attendants. Such sessions can be attacked and must be secured by suitable tools, e.g. for detection of morphing attacks and transmission of deep fakes. The workshop examines the current status of this approach and discusses future steps.

Program

FAKE-ID: A BMBF funded research project
(Dr. Klaus Hermann, Maurer Electronics GmbH)

Being able to unambiguously check or prove the identity (ID) of a person and the authenticity of images and videos on the Internet in every situation becomes more complex and technically demanding due to the merging of real and virtual processes.

The aim of FAKE-ID is to identify and classify possible attacks and forgeries of images and video data streams using different types of "fake IDs" at the borderline between physical and virtual processes in image and video-based authentication processes, and to develop algorithms for their recognition. In order to be able to recognize attack scenarios and forgeries and evaluate them in a well-founded manner, the project generates new insights into the diverse motivations and backgrounds (from fraud to the manipulation of political decision-making processes) for the falsification of images and video data streams and how police authorities and the judiciary deal with such attacks and forgeries.

Features and properties of real as well as forged identities (ID) will be formally described and their detection and exclusion is learned using weak artificial intelligence (AI) in order to make authenticity indicators and suspicions for the presence of forged ID (deep fakes) identifiable in images and videos. The project results will make it possible to communicate these indications and suspicions to the ultimately decisive human personnel using a risk and suspicion map as a basis for decision-making in order to verify and understand this. The associated process should be able to be embedded in image and video-based proceedings and thus be made available in a generic manner that is adapted to the needs of judicial proceedings.

ANANAS: A BMBF funded research project

(Dr. Christian Krätzer, Otto-von-Guericke University Magdeburg)

Between June 1st, 2016 and May 31st, 2020, the research project ANANAS (English: "Anomaly detection to prevent attacks on facial image-based authentication systems", German: "Anomalie-Erkennung zur Verhinderung von Angriffen auf gesichtsbildbasierte Authentifikationssysteme") performed research on one recent and significant threat for face image based authentication scenarios: the threat imposed by face morphing attacks. In these attacks, face images of two (or more) persons are blended together (morphed) to create an output image that is perceptually as well as biometrically similar to all inputs. If such a morph would be used for an application for a new document (such as a passport), this document could successfully be used by multiple persons. The attack itself was first described as a theoretical threat to biometric systems in 2009, was published upon with a detailed step by step description in 2014 and is known to have been exploited in 2018 to successfully apply for at least one German passport.

The ANANAS project, funded by the German Federal Ministry of Education and Research, made significant progress in addressing this threat. The outcomes of this very successful project are including a large number of scientific publications, concepts for face morphing attack detectors, evaluation datasets, multiple demonstrators (in hard- and software), input to standardization as well as two patent applications.

The Joint project is summarized with the overall research performed in the field of morph detection. Exemplary scenarios for document misuse, as usage of stolen passport, passport forgeries and face morphing attacks (FMA) are discussed. FMA are targeting the template space of a biometric application. The results in modelling, detectors, demonstrator(s) and standardization are presented. The work of three partner institutions include more than ten researched, implemented detection approaches, which are integrated into the ANANAS demonstrator(s).

Security and trust in mobile AI applications

(Dr. Maxim Schnjakin, Bundesdruckerei GmbH)

Artificial Intelligence (AI) methods, such as Machine Learning (ML), e.g. Deep Learning (DL), are already being used in a variety of continuously growing heterogeneous mobile and embedded platforms. Due to their heterogeneity, it is difficult to protect such platforms against cyber-attacks and to prevent the manipulation of ML/DL models. This can have life-threatening consequences, e.g. in the field of autonomous driving. But even in less critical environments, such as on smartphones, a targeted attack on AI systems can have far-reaching consequences. For example, targeted manipulation of the input can bypass the biometric face recognition of the authentication process.

The planned project aims to make such attacks on AI systems in mobile and embedded contexts more difficult or to prevent them completely. Proven mechanisms of trusted computing will be used for this purpose. With the help of scientific methods, it will be investigated whether such technologies offer effective and reliable protection in the assumed attack scenarios. In addition to hardware-based methods, software methods for monitoring and integrity preservation will also be developed. The goal of the project is to integrate AI systems into mobile and embedded devices (e.g. smartphone, IoT surveillance camera, edge server) in a secure and easy-to-use way. Two different prototypes will be developed to evaluate the effectiveness and efficiency of the developed approach.

Explaining deep neural network decisions for secure verification

(Prof. Dr. Peter Eisert, Fraunhofer Heinrich Hertz Institute)

Recent advances in deep learning have fostered the use of deep neural networks for almost arbitrary applications, including biometrics and forensics. However, the improved performance and accuracy compared to classical approaches is usually traded against results that are more difficult to interpret. Deep neural networks typically are black box systems, which are often critical in security and safety related applications. Thus, it would desirable to know, why a neural network has come to a particular decision in order to verify a decision or enhance the performance of the system in case of mis-classifications. In this talk, recent activities in the detection of face morphing, deep fake and presentation attack detection were demonstrated, highlighting the performance of black box deep neural network architectures. In order to make network decisions more transparent, methods like layer wise relevance propagation (LRP) were used to find weaknesses in the detectors and to enhance them towards robustness against unseen new attacks. Focused layer wise relevance propagation (FLRP) can also help non-technical experts (e.g. border guards) to understand the black box systems. For all the attack detection methods, it is shown that not only plausibility of decisions can be determined but also generality and attack detection performance can be improved.

Security technology and transparency

(Prof. Dr. Hartmut Aden, Dr. Jan Fährmann, Prof. Dr. Sabrina Schönrock, Berlin School of Economics and Law)

New security technologies can have a major impact on the use of fundamental rights such as privacy. Therefore the development and use of new security technologies have to be held accountable, based on risk assessments and other standardized procedures. This also includes the often neglected perspectives of the individuals affected by security technologies: Do they understand (and accept) what security agencies do with their personal data? The procedural justice theory that we apply in this context claims that transparency facilitates the acceptance of policing in general. This can also be applied to the use of security technologies.

Security technologies such as the detection of deep fakes– as they are to be developed in the FAKE-ID project – are based on AI. Therefore they have to comply with specific transparency requirements with respect to the development and use of AI-based tools. This includes the explainability of results and transparency of automated suggestions for decisions from the perspective of the final (human) decision-maker and for those affected by the decisions. The draft EU regulation presented by the European Commission in April 2021 (COM (2021)206 final) is expected to set the standards for the use of AI in Europe and beyond if it enters into force.

From morph attack detection to DeepFake detection using hand-crafted features

(Prof. Dr. Jana Dittmann, Otto-von-Guericke University Magdeburg)

A summary of the work in Siegel, D.; Kraetzer, C.; Seidlitz, S.; Dittmann, J. Media Forensics Considerations on DeepFake Detection with Hand-Crafted Features. J. Imaging 2021, 7, 108. (https://www.mdpi.com/2313-433X/7/7/108) is presented. Features hand-crafted by domain experts as an alternative approach of DeepFake for Video are discussed and the main advantage of interpretability for plausibility validation for decisions are presented. Three sets of hand-crafted features and three different fusion strategies and evaluation results of DeepFake detection are summarized. Further the forensics data model from Kiltz, S. Data-Centric Examination Approach (DCEA) for a qualitative determination of error, loss and uncertainty in digital and digitised forensics.

PhD thesis, Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, 2020 is used to show what data in the signal and pattern recognition pipeline is used during detection to support transparency and explainability of results achieved. For expert assessment, the system causability scale, work from Andreas Holzinger, André Carrington, and Heimo Müller. 2020. Measuring the quality of explanations: the system causability scale (SCS). KI-Künstliche Intelligenz (2020), 1–6, is explored. In future, the requirements from BSI: AI Cloud Service Compliance Criteria Catalogue (AIC4) (https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/AIC4/aic4) are planned to be included as well.

Liveness detection and deep fakes status quo
(Ann-Kathrin Freiberg, BioID GmbH)

For deriving reliable credentials through digital identity verification, ongoing development is required to cope with new challenges and attacks, e.g. deep fakes. Face liveness detection is an anti-spoofing method for facial biometrics. Scientifically, it is called presentation attack detection (PAD). The core function of a PAD mechanism is to determine whether a biometric feature (e.g. a picture), was captured from a live person. ISO/IEC 30107-3 compliant liveness detection prevents biometric fraud through printed photos, cutouts, prints on cloth, 3D paper masks, videos on displays, video projections and more. Deep fakes presented at the sensor level (e.g. on displays) can be rejected through the same methods, e.g. texture analysis and artificial intelligence. For counter fighting attacks at the application level, namely through the injection of modified camera streams, challenge-response mechanisms can be utilized to reject prerecorded videos/deep fakes. But, deep fake modification can take place in real-time, projecting a face onto a live moving face, which could overcome even challenge-response technology.

Thus, for both, the agent-based as well as the fully remote identity verification processes, deep fakes (especially injected via virtual cameras) are becoming a growing challenge. Ensuring secure applications which reject virtual cameras and modified video streams as input prevents deep fake attacks during automated selfie verification. At the same time, research needs to be pursued for detecting deep fakes in the photo/video material directly. This is the aim of the consortium created for FAKE-ID.

Fake identity detection in speech data
(Nicolas Müller, Fraunhofer AISEC)

With the advances in artificial intelligence, the so-called 'deep fake' technology is also gaining strength, making it possible to ‘digitally clone’ a target person. Specifically, this technology can be used to put arbitrary words in the mouth of a natural person (see, for example, this clone of Angela Merkel https://www.youtube.com/watch?v=MZTF0eAALmE). Although this technology has some benign applications (movie and TV industry, grief counselling, Human-PC interfaces) the potential for abuse is enormous. This talk will show under what conditions targets can be cloned, and how AI-based defense systems attempt to detect such forgeries.

Panel discussion
The panelists,

- Prof. Dr. Hartmut Aden (Berlin School of Economics and Law)
- Prof. Dr. Jana Dittmann (Otto-von-Guericke University Magdeburg)
- Prof. Dr. Peter Eisert (Fraunhofer Heinrich Hertz Institute)
- Dr. Eleanor Hobley (Central Office for Information Technology in the Security Sector ZITiS)

- Kornelia Nehse (Berlin Police, Criminal Investigation Department)
- Dr. Andreas Wolf (Bundesdruckerei GmbH)

Discussed many topics related to remote identification and responsible use of AI technologies. They spoke on questions like:

- In Germany, we have an identity management system in place, which is based on the German ID Card (Personalausweis), supported by residence permits. Other EU Member States will deploy eIDAS compliant systems as well. This would allow for a PKI based identity handling in the Internet. Do we still need fake detection capabilities?
- Fake detection in video streams and the like is a tool for supporting secure remote authentication. This tool makes use of face or voice data, that is, of personal information. This data might contain excessive information, revealing, e.g., health information. How do we protect such data while ensuring a secure authentication?
- Morphing is a potential threat to portrait based ID documents. Assuming all face enrolment is done in a controlled environment, e.g., in a municipality office, would we still need morphing detection capabilities? Or, vice versa, would good morphing and presentation attack detection mechanisms allow for much more relaxed enrolment conditions, maintaining a high level of security and data quality as well?
- AI technologies might lead to non-transparent decisions which could be considered problematic in a democratic society. How can we ensure that AI remains a tool to the decision maker increasing their capabilities to make good decisions, and not something that leads to discriminating, biased and finally unfair results?