Cost Action CA-17124 DigForASP:
A European cooperative Action for AI applications in Police and Digital Investigations
An Exploration of Investigation Complexity

Raffaele Olivieri[1], Jesús Medina[2], Stefania Costantini[3], Francesca Alessandra Lisi[4],

[1] WG1 Leader - Arma dei Carabinieri - Ra.C.I.S. - `raffaele.olivieri@gmail.com` (Italy)
[4] Chair - University of Cadiz - `jesus.medina@uca.es` (Spain)
[2] vice-Chair - University of L'Aquila - `stefania.costantini@univaq.it` (Italy)
[3] Science Communication Manager - University of Bari - `francesca.lisi@uniba.it` (Italy)

Abstract

In the frame of Police Investigations, in particular to Digital Investigations and Digital Forensics cases, data collection on "crime scene" needs further elaboration for the contextualization in the real case. The Evidence Analysis phase has the aim to provide objective data and suitable elaboration of these data can help the Investigators in the formulation of possible investigative hypotheses, which could later be presented as proofs of evidence in courts. Investigations with a high amount of heterogeneous data represent a huge problem for the human mind in the search for events, connections, facts or demonstrate alternative solutions. However, many investigative problems can be formalized and expressed with a mathematical approach and solved with reasonable efficiency with Artificial Intelligence and Automatic Reasoning. COST Action CA17124, called DigForASP (*DIGITAL FORensics: analysis tests through intelligent systems and practices*), financed by the European Union with the funds for "European cooperation in science and technology, Horizon 2020", was born for the exploration, study the delicate issue of the application of Artificial Intelligence and Automated Reasoning to the investigative world, through the creation of a multidisciplinary scientific network. DigForASP, with activities in the period September 2018 - September 2022, has aims to help the human operator (Law Enforcements, Lawyers, Public Prosecutors, Judges, social scientists, criminologists) in the analysis of investigative data as well as the formulation of hypotheses for the resolution of complex cases, through Artificial Intelligence techniques available to guarantee ethic, reliability and verifiability.

Introduction

"Digital Forensics" is a part of the Criminalistics Sciences which deals with digital evidence recovery and exploitation in the solution of criminal cases through the application of scientific principles. There are several and increasingly sophisticated methods for collecting digital evidence. As a matter of fact, the evolution of technology continuously pushes such kind of methods. Rough evidence must however be used to elicit hypotheses concerning events, actions and facts (or sequences of them) with the goal to obtain evidence to present in court. Evidence analysis involves examining fragmented incomplete knowledge, and reconstructing and aggregating complex scenarios involving time, uncertainty, causality, and alternative possibilities. No established methodology exists today for digital evidence analysis. The Scientific Investigation experts usually proceed by means of their experience and intuition. The Challenge of the proposed COST Action consists in creating a Network for exploring the potential of the application of Artificial Intelligence and Automated Reasoning in the Digital Forensics field, and creating synergies between these fields. Specifically, the challenge is to address the Evidence Analysis phase, where evidence about possible crimes and crimes perpetrators collected from various electronic devices (by means of specialized software, and according to specific regulations) must be exploited so as to reconstruct possible events, event sequences and scenarios related to a crime. Evidence Analysis results are then made available to law enforcement, investigators, public prosecutors, lawyers and judges: it is therefore crucial that the adopted techniques guarantee reliability and verifiability, and that their result can be explained to the human actors.

Structure and Activities

DigForASP has more than 200 participants (of which 40 % women) from 35 of European Union countries, 19 ICT countries ("Inclusiveness Target Countries", and Ukraine as NNC (Near-Neighbour Country.

The Action Coordinator is **Prof. Jesús Medina** of the University of Cadiz, the Vice-Coordinator is **Prof. Stefania Costantini** from the University of L'Aquila and the Science Communication Manager is **Prof. Francesca Alessandra Lisi** of the University of Bari.

The Action is based on ideas developed in the Doctoral Thesis by **Raffaele Olivieri**, a Warrant Office of Italian Carabinieri (supervisor Prof. Stefania Costantini) and provides the following scientific packages.

WG	Name	Leaders	Country
WG1:	Digital Forensics requirement analysis	Dr. Raffaele Olivieri	Italy
WG2:	Research on applications of AI/AR to DF	Prof. Alessandra Mileo	Ireland
WG3:	Prototypes and Platforms	Prof. Pedro Cabalar	Spain
WG4:	Benchmarks based on real cases	Prof. Viviana Mascardi	Italy
WG5:	Platform integration and multi-dimensional enviroments	Prof. Esra Erdem	Turkey
WG6:	Meetings, workshops and conferences	Prof. Andreja Tepavcecic	Serbia
WG7:	Training, education and dissemination activities	Dr. Aleksandra Dedinec	Macedonia
WG8:	Short-Term Scientific Missions and internal organizzation	Prof. Manuel Ojeda Aciego	Spain

The activities promoted by the Action are from time to time published on the website: https://digforasp.uca.es/. They allow computer scientists to understand the problems main and the open problems of DF, and in particular of the phase of as Evidence Analysis ". They will also help promote the exploitation of AI to deal in an innovative way, effective and adaptive key issues in the DF domain. The Action partners have identified applicable techniques, and they will have to suggest guidelines for creation and development of new techniques and methods aimed at advancing the state of the art in both areas. This will strengthen European research and the capacity for innovation in these areas.

Main Objectives

1. **WG1 Digital Forensics requirement analysis: Leader Dr. Raffaele OLIVIERI**
Identify problems, in Digital Forensics sectors and in its analysis processes, that involve the examination of incomplete or fragmented knowledge, or which include complex scenarios about the evolution of time, causality, uncertainty, doubts, randomness and the existence of alternative solutions:
   - O1: Raising awareness among researchers in AI and Automated Reasoning of the main issues and problems in the daily work of digital forensics scientists and the technical, legal and criminological aspects they involve.

2. **WG2 Research on applications of AI/AR to DF: Leader Dr. Alessandra MILEO**
Promote and facilitate understanding of the Digital Forensics domain through targeted multidisciplinary activities, including characterisation of real use-cases as well as analysis of available digitized datasets. Design computational methods that are able to take into account the diversity of multiple evidence in terms of their granularity, legal remit, quality, relevance, timeliness, sensitivity. Develop techniques to ensure that applied automated reasoning methods include human understandable explanations.
   - O2: Identify a list of realistically applicable AI and AR techniques, focusing on Computational Logic for verifiability and justifiability reasons (essential in a legal environment);
   - O3: Promote R&D activities for integrating such techniques and defining new methods and tools;
   - OC1: Increase awareness among DF experts of new methods and techniques and their applicability.

3. **WG3 Prototypes and Platforms: Leader Prof. Pedro CABALAR**
Development of software prototypes that apply techniques of AI/AR (as those spotted in WG2) to different practical scenarios in DF (extracted from WG1). These prototypes may combine several existing (or adapted) platforms for AR/AI and will be tested on benchmarks, either synthentically generated or using anonymized data from real cases:
   - O3: Promote R&D activities for integrating such techniques and defining new methods and tools.
   - O5: Collect and organise anonymised data extracted from real cases provided by the DF Action partners, and devising suitable benchmarks concerning the solution of such cases.
   - O6: Put into practice the new methods developed in the Action on the collected data thus providing a foundation for future practical tools.

4. **WG4 Benchmarks based on real cases: Leader Prof. Viviana MASCARDI**
Development of a suite of benchmarks for testing the prototypes developed in WG3. The scenarios addressed by the benchmarks and the languages for representing them will come from WG1. With the help of WG3 and WG5, and with the active involvement of domain experts, WG4 will identify qualitative and quantitative measures suitable for assessing the prototypes performances, and will validate and verify such performances by running the test cases in the benchmark suite:
   - O3: Promote R&D activities for integrating such techniques and defining new methods and tools.
   - O5: Collect and organise anonymised data extracted from real cases provided by the DF Action partners, and devising suitable benchmarks concerning the solution of such cases.
   - OC1: Increase awareness among DF experts of new methods and techniques and their applicability.

5. **WG5 Platform integration and multi-dimensional environments: Leader Prof. Esra ERDEM**
Development of methods/software to integrate a variety of AI/AR/DF platforms (WG3) and multidimensional data/knowledge resources (WG4), which are developed for different KRR tasks/applications related DF (WG2), and to illustrate their applications on practical scenarios(WG1):
   - O5: Collect and organise anonymised data extracted from real cases provided by the DF Action partners, and devising suitable benchmarks concerning the solution of such cases.
   - O6: Put into practice the new methods developed in the Action on the collected data thus providing a foundation for future practical tools.

6. **WG6 Meetings, workshops and conferences: Leader Prof. Andreja TEPAVCEVIC**
Under WG6 various types of meetings, workshop and conferences will be organized in order to share results within the research areas related to the project. Moreover, the project will organize special sessions and project members will participate at the different meetings in digital forensics and related areas organized by other colleagues in the field. In this way project members would meet best experts in theory and practice in the field and as a synergy, new methods and techniques will be developed:
   - O4: Foster and coordinate related R&D activities of the partners.
   - O5: Put into practice the new methods developed in the Action on the collected data thus providing a foundation for future practical tools.
   - OC1: Increase awareness among DF experts of new methods and techniques and their applicability.

7. **WG7 Training, education and dissemination activities: Leader Dr. Aleksandra DEDINEC**
Produce and maintain the Action Website as well as oversee Social Media presence, collect and study the different proposals to organize Training Schools and dissemination activities and publications in particular propose Training Schools and dissemination activities and publications to the MC. Study the negative impact of the given results and advances before the publication and Guarantee that the Action preserves the societal, ethical and legal aspects, and Responsible Research and Innovation.
   - O4: Foster and coordinate related R&D activities of the partners.
   - OC1: Increase awareness among DF experts of new methods and techniques and their applicability.
   - OC3: Disseminate the Action results to stakeholders, not only Government Agencies and Institutions but also private companies working in the DF field.
   - OC4: Implement training schemes for the parties that are potentially involved.

8. **WG8 Short-Term Scientific Missions: Leader Prof. Manuel OJEDA-ACIEGO**
Foster and coordinate R&D activities of the partners. Particularly, increase awareness among DF experts of new methods and techniques and their applicability, and assist them in the practical experimentation of the new methods, and vice versa.
   - O4: Foster and coordinate related R&D activities of the partners.
   - OC1: Increase awareness among DF experts of new methods and techniques and their applicability.
   - OC2: Assist DF experts in the practical experimentation of the new methods.

Conclusions

It has been possible to highlight the usefulness of the Answer Set Programming (ASP, introduced in [5, 6] and illustrated in [1]) to examine investigative scenarios. In fact, an investigation often seems to admit only non-linear or even chaotic solutions. However, after a methodical analysis and decomposition into the elementary components, many cases can be represented through a rigorous mathematical approach. The form that cases (or fragments of cases) take is not infrequently typical of known optimization problems, belonging to the complexity classes P or NP, which can be expressed, and can often be resolved with reasonable efficiency, using ASP. This is illustrated in the seminal work [2] and in the subsequent works [1, 3, 4].

References

[1] Gerhard Brewka, Thomas Eiter, and Miroslaw Truszczynski (eds.). Answer set programming: Special issue. *AI Magazine*, 37(3), 2016.

[2] Stefania Costantini, Giovanni De Gasperis, and Raffaele Olivieri. Digital forensics evidence analysis: An answer set programming approach for generating investigation hypotheses. In *Logic Programming and Nonmonotonic Reasoning - 13th International Conference, LPNMR 2015, Proceedings*, volume 9345 of *Lecture Notes in Computer Science*, pages 228–241. Springer, 2015. Long version in CEUR Workshop Proceedings of CILC 2015, 30th Italian Conference of Computational Logic.

[3] Stefania Costantini, Giovanni De Gasperis, and Raffaele Olivieri. Digital forensics and investigations meet artificial intelligence. *Annals of Mathematics and Artificial Intelligence*, 2019. pp 193-229 Volume 86/2019 ISSN: 1012-2443 (Print) 1573-7470 (Online), Springer International Publishing.

[4] Stefania Costantini and Raffaele Olivieri. Digital forensics meets complexity theory and artificial intelligence. *Digital Forensics Magazine*, 2019. May 2019 Issue 39: pp- 50-55.

[5] Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In Robert Kowalski and Kenneth Bowen, editors, *Proc. of the 5th Intl. Conf. and Symposium on Logic Programming*, pages 1070–1080. MIT Press, 1988.

[6] Michael Gelfond and Vladimir Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9:365–385, 1991.