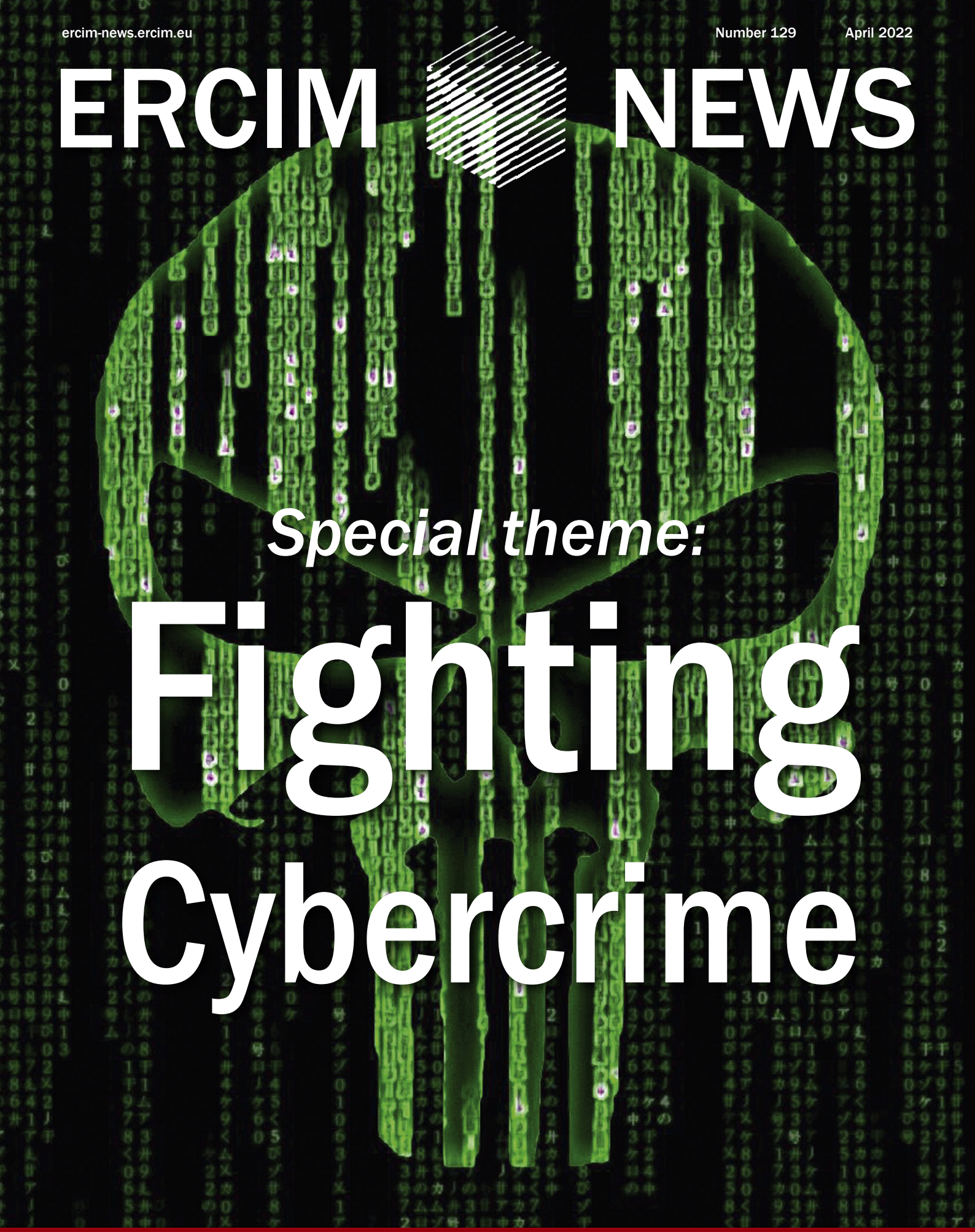
ercim-news.ercim.eu Number 129 April 2022

ERCIM NEWS



Special theme:

Fighting Cybercrime

Also in this issue

Research and Innovation:

Artificial Intelligence Enabled Distributed Edge Computing for Internet of Things

Editorial Information

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online ,at https://ercim-news@ercim.eu.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
+33 4 9238 5010, contact@ercim.eu
Director: Philipp Hoschka, ISSN 0926-4981

Contributions
Contributions should be submitted to the local editor of your country

Advertising
For current advertising rates and conditions, see https://ercim-news.ercim.eu/ or contact peter.kunz@ercim.eu

ERCIM News online edition: *https://ercim-news.ercim.eu/*

Next issue:
July 2022: Assistive and Inclusive Technologies

Subscription
Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu

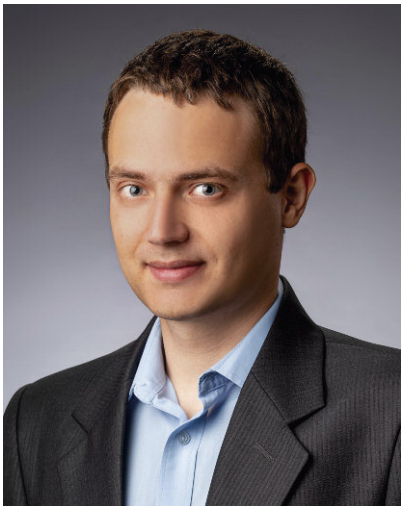
JOINT ERCIM ACTIONS

SPECIAL THEME

The special theme "Fighting Cybercrime" has been coordinated by the guest editors Florian Skopik (AIT Austrian Institute of Technology) and Kyriakos Stefanidis (ISI)

Tomasz Kociumaka wins the 2021 ERCIM Cor Baayen Young Researcher Award

The ERCIM Cor Baayen Award selection committee has unanimously selected Tomasz Kociumaka as the winner of the competition for the 2021 ERCIM Cor Baayen Young Researcher Award. Tomasz Kociumaka was nominated by the University of Warsaw, which awarded him a PhD in 2019. Tomasz then worked at Bar-Ilan University, Israel, and he is currently a postdoctoral researcher at the University of California, Berkeley, USA.



Tomasz Kociumaka.

In his short career as a young researcher, Tomasz has already published at the world's most important conferences in theoretical computer science such as STOC, FOCS, and SODA. His research concerns fundamental issues in computer science and solves conjectures and problems that have remained unsolved for many years. According to Google Scholar, his work has been cited more than 1400 times, which is remarkable for a researcher working on theoretical aspects of computer science under three years after completing his PhD.

At a high level, Tomasz's research interests lie in designing efficient algorithms for processing strings (arbitrary sequences of characters) and exploring the underlying combinatorial problems. This is a classic research area with a number of problems central to computer science, such as pattern matching or compression. But this also means that, after half a century of research, new developments that significantly contribute to these very natural problems are extremely hard to come by, rare, and only accessible to elite researchers. There is no doubt that Tomasz has found his place in this category, with multiple groundbreaking results achieved in the last six years.

It is very likely that Tomasz's work will lead to algorithms that are used in practice, as algorithms for processing large and often compressed data are ubiquitous in many applications. For example, a data structure proposed by Kempa and Kociumaka for processing the Longest Common Extensions queries (STOC'19) was implemented with a team of algorithm engineers at TU Dortmund (Dinklage, Fischer, Herlez, Kociumaka, Kurpicz), and a resulting paper (ESA'20) shows the practicality of the approach with appropriate fine-tuning and simplifying heuristics.

In summary, Tomasz Kociumaka clearly stands out from his generation of computer scientists. In his short career, he has already (co)-authored more than 90 publications, many of which have been presented at high-level conferences. He was a member of the programme committee of nine conferences and also serves the scientific community as an organiser of conferences, workshops, and programming competitions for high school students. The selection committee for the ERCIM Cor Baayen Young Researcher Award is proud to present the 2021 award to Tomasz Kociumaka, one of the most promising young researchers in the field of computer science or applied mathematics.

Cor Baayen Award 2021

Winner:

- Tomasz Kociumaka (IEOR Department, University of California, Berkeley, USA), nominated by Łukasz Kowalik (University of Warsaw)

Honorary mention:

- Clara Stegehuis (Twente University, Department of Mathematics, Electrical Engineering and Computer Science), nominated by Ton de Kok (CWI);
- Lucia Vadicamo (Institute for the Science and Technologies of Information, Italian National Council of Research - ISTI-CNR), nominated by Giuseppe Amato (ISTI-CNR).

Finalists:

- Giulio Ermanno Pibiri (CNR) nominated by Raffaele Perego (CNR);
- Christopher Krauß (Fraunhofer Institute for Open Communication Systems - FOKUS) nominated by Dieter Fellner (Fraunhofer ICT Group);
- Johannes Späth (Fraunhofer Institute for Mechatronic Systems Design IEM) nominated by Dieter Fellner (Fraunhofer ICT Group);
- Elena Gaburro (Inria Bordeaux Sud-Ouest) nominated by Mario Ricchiuto (Inria);
- Thomas Debris-Alazard (Inria) nominated by Jean-Yves Berthou (Inria);
- Chiara Sironi (Maastricht University) nominated by Ton de Kok (CWI).

Evaluation Committee:

The Evaluation Committee was composed of Monica Divitini (NTNU, chair of the ERCIM Human Capital Task Group), Thierry Priol (Inria), Fabrizio Sebastiani (ISTI-CNR) and Jerzy Tiuryn (UWAW). The decision was unanimous.

More information about the ERCIM Cor Baayen Young Researcher Award:
https://www.ercim.eu/human-capital/cor-baayen-award

ERCIM “Alain Bensoussan” Fellowship Programme

The ERCIM PhD Fellowship Programme has been established as one of the premier activities of ERCIM. The programme is open to young researchers from all over the world. It focuses on a broad range of fields in Computer Science and Applied Mathematics.

The fellowship scheme also helps young scientists to improve their knowledge of European research structures and networks and to gain more insight into the working conditions of leading European research institutions. The fellowships are of 12 months duration (with a possible extension), spent in one of the ERCIM member institutes. Fellows can apply for second year in a different institute.

Why to apply for an ERCIM Fellowship?

The Fellowship Programme enables bright young scientists to work on a challenging problem as fellows of leading European research centers.An ERCIM fellowship helps widen and intensify the network of personal relations among scientists.

The programme offers the opportunity to ERCIM fellows:

- to work with internationally recognized experts;
- to improve knowledge about European research structures and networks;
- to become familiarized with working conditions in European research centres;
- to promote cross-fertilization and cooperation, through the fellowships, between research groups working in similar areas in different laboratories.

Conditions

Candidates must:

- have obtained a PhD degree during the last eight years (prior to the year of the application deadline) or be in the last year of the thesis work;
- be fluent in English;
- have completed their PhD before starting the grant.

The fellows are appointed either by a stipend (an agreement for a research training programme) or a working contract. The type of contract and the monthly allowance/salary depends on the hosting institute.

Application deadlines

Deadlines for applications are currently 30 April and 30 September each year.

Since its inception in 1991, over 750 fellows have passed through the programme. In 2021, 26 young scientists commenced an ERCIM PhD fellowship and 54 fellows have been hosted during the year. Since 2005, the Fellowship Programme is named in honour of Alain Bensoussan, former president of Inria, one of the three ERCIM founding institutes.

http://fellowship.ercim.eu

ERCIM Working Group
Dependable Embedded Systems

SAFECOMP 2022 and the DECSoS 2022 Workshop

Munich and online 6-9 September 2022

Invitation for participation

SafeComp has contributed since 1979 to the progress of the state-of-the-art in dependable application of computers in safety-related and safety-critical systems. SafeComp is an annual event covering the state-of-the-art, experience and new trends in the areas of safety, security and reliability of critical computer applications. SafeComp provides ample opportunity to exchange insights and experience on emerging methods, ap-proaches and practical solutions. It is a single-track conference allowing easy networking. SAFECOMP 2022 will take place on 6-9 September 2022 at Fraunhofer AISEC and Galileo Science Technolgie Park in Munich Garching, a few subway stops from Munich city center.

Call for papers

The DECSoS Workshop (17th International Workshop on Dependable Smart Embedded Cyber-Physical Systems and Systems-of-System) was created by the ERCIM Dependable Embedded Systems Working Group and is still continuing successfully. Topics cover a large scope in the context of dependable, trustworthy systems. It is collocated with SAFECOMP as one of eight workshops on September 6th, at the same venue, with physical and online participation (hybrid). The call for papers for DECSoS is still open, the deadline is May 6th, 2022. Workshop papers are reviewed by at least three independent reviewers. Accepted papers will be published by Springer Nature in the LNCS series as “SAFECOMP 2022 Workshop Proceedings”.

For details, see
https://safecomp22.iks.fraunhofer.de/

Workshops are listed under the tag “Workshops”.

Please contact:
Erwin Schoitsch, AIT Austrian Institute of Technology),
erwin.schoitsch@ait.ac.at

Amund Skavhaug (NTNU, Trondheim, Norway), amund.skavhaug@ntnu.no

Introduction to the special theme

Fighting Cybercrime

by Florian Skopik (AIT Austrian Institute of Technology) and Kyriakos Stefanidis (ISI)

Cybercrime has grown to a profitable multi-billion-dollar business. The number of reported criminal offences is continuously rising every year. The reasons for this development are the ever-increasing dependency on IT technology for almost every business, the opportunity for attackers to operate in the dark, and the continuously growing attack surface. With the adoption of new computing paradigms, such as cloud computing and the Internet of Things, not only new opportunities for legitimate businesses arise, but also new ways for criminals to make profit or to attack and de-stabilise a country's economy or society. In recent years, we have witnessed the rise of ransomware attacks on a large scale, Distributed Denial of Service (DDoS) attacks with high volumes that have never been observed before, and data leaks that massively harmed global businesses. Besides stealing business-critical data, harming or blackmailing individuals or organisations, large-scale attacks on critical infrastructures of a region or nation state have become a severe threat. Some examples are the recent attacks on the US-East-Coast Colonial Pipeline, one of the largest US pipeline operators, and the use of cybersecurity attacks on Ukrainian infrastructures. Finally, the ever-growing de-stabilising disinformation campaigns and cyberwar practices in general are increasingly shaping social and political conflicts. Thousands of high-impact attacks have already demonstrated the vulnerability of complex interconnected systems.

The predicted developments for the coming years are worrying. The use of ransomware has picked up pace and many security companies expect a rapid rise regarding their variations and frequency. Attacks on operational technology (OT) are evolving from rather simple process disruption, such as shutting down a plant or factory, to compromising the integrity of industrial environments with intent to cause physical harm. Further, the pandemic led to remote working on a large scale, which changed the IT environment tremendously. Home devices that employees use to access office networks are usually not subject to the same security restrictions as corporate devices. This complicates efforts to control and monitor employees' digital behavior, applications, and data outside the carefully monitored business environment. Nearly half of the organisations moved business-critical functions to the cloud as a direct result of the pandemic. While this migration to a professionally managed environment will increase security for many customers, problems arise at the interfaces between a company's own systems and highly virtualised remote data centers. This increases the attack surface in many cases. Additionally, such a migration changes common security workflows. For instance, detecting and preventing malicious activity in a multi-tenancy cloud tremendously differs from doing the same in a traditional on-premises setup. Last, but not least, we currently witness the effects of geopolitical events on cybercrime activities. State actors may actively encourage criminals to carry out cybercrime activities, or launch attacks themselves, because they are cheap, reliable, scalable, and hard to attribute. Coping with these situations seems overwhelming, however we are not defenseless. Research on new concepts, methodologies, technologies, and tools is vital to protect our digital infrastructures from adversarial activities.

In the fight against cybercrime there are several European agencies, organisations and initiatives that help the law enforcement agencies (LEAs) both at the level of operations and also technical capacity. Europol's European Cybercrime Center [L1], since 2013 provides operational and technical support to LEAs, provides training and capacity building to the relevant authorities, and helps with the collaboration of LEAs with the other cyber communities, bodies and agencies such as ENISA, CERT-EU, etc. The European Union Agency for Cybersecurity (ENISA), since 2004, was founded to enhance the capability of the Member States to prevent, address and to respond to network and information security problems. It also supports the cooperation between the national Computer Security Incident Response Teams (CSIRTs)[L2] via the CSIRT Network. In this special theme, we have included articles from European CSIRTs as well as LEAs that describe their novel approaches in incident response and fighting cybercrime in general.

The articles in this special theme cover five broad areas of cybersecurity and more specifically cybercrime. Therefore, we divided them into five different groups. The first group includes general law enforcement investigation processes as well as the evolution of cybercrime from a technological and business point of view. Then we have several articles that cover technical approaches to detect and counter cybercrime activities. A good proportion of them use Artificial Intelligence (AI) and Machine Learning (ML) as their enabling technologies. Besides the technical approaches, we have a group of articles that focus on novel incident response and threat intelligence processes. We conclude with a group of articles that present educational and awareness-raising initiatives.

Evolutions in cybercrime and law enforcement investigation processes

IIn our first group of papers, we deal with the operational and business side of cybercrime and law enforcement. Markatos (page 8) discusses the technical drivers of cybercrime and how it becomes more organised by using established business models such as *-as-a-service or alternative forms of wealth transfer, such as cryptocurrencies, to scale-up their business. On the other hand, from the law enforcement side, King et al. (page 9) presents a novel approach to counter terrorist financing that instead of the common "follow the money" strategy, proposes a "follow the actor" approach for financial investigations. Tsakalidis et al. page 11) also shows how financial investigators use a

BPMN-based investigation process for copyright-related cybercrime offences.

Machine Learning and AI to detect cybercrime activities

Many fields of computer science use AI and ML as enabling technologies for their purposes. Security and privacy follow the same trend. Ferreira et al. (page 13) presents an SVM-based digital forensics tool for the detection of deep fake images. Part of their work is also a comprehensive dataset of images and videos suitable for digital forensics. Mayer (page 14) proposes a privacy-preserving anomaly-detection method that is suitable for confidential data analysis by third parties. The approach is based mainly on collaborative learning and synthetic data. Han et al. (page 16) focus on ML in resource-constrained edge devices. Similarly, Buttyán et al. page 17) focus on malware analysis specifically for resource-constrained devices (IoT) and present their ML-based approach. Again, on the topic of malware analysis, Iadarola et al. (page 19) discuss the trustworthiness and explainability of deep learning techniques and how/when a security analyst can trust the predictions of those. Ceolin (page 21) is also dealing with AI trustworthiness but focuses on the field of information quality and the spread of mis/disinformation online.

Further technical approaches to counter cybercrime

Although many novel approaches harness ML in different ways to become more effective, we must not neglect further technical solutions. Kern et al. (page 22) ask the question of what log and network data need to be collected in the first place to spot adversarial activities at all, e.g., by the means of ML-based intrusion detection systems. Landauer et al. (page 24) discuss an approach to flexibly create testbeds in a model-driven manner. These testbeds allow to benchmark and validate the effectiveness of intrusion detection systems before they are deployed in productive environments. Folino et al. (page 25) introduce an approach based on the popular ELK framework to collect data that reflect digital user behaviour footprints and to subsequently detect anomalies in these data.

Timestamps of files are a great source of information for forensic investigations of cybercrime activities. Luh et al. (page 27) discuss tools to detect timestamp forgery, which is essential for reliable results of investigations. Often malware is being used in criminal activities. The analysis of malware to better understand their mode of operation and potential harm is a time-consuming task. Thus, Kochberger et al. (page 28) introduce a meta-framework for automating static malware analysis.

Effective security processes, incident response and threat intelligence

Technical means to cybersecurity are essential, but effective standards, processes and procedures are of at least equal importance. A vital component of cybersecurity is establishing situational awareness. Knowing and understanding emerging trends helps to detect changes in the threat landscape and may have a considerable impact on security governance. Kohlrausch (page 30) demonstrates the augmentation of security metrics with stochastic models to keep track of trends, which is a cornerstone of justified decision making. A prerequisite of situational awareness is the collection of information in the first place. Sharing of incident handling information, the automation of incident response processes, as well as the relationship between these two topics, to assist human operators in their work is therefore at the centre of Nitz et al. (page 31) and their European SAPPAN project.

Furthermore, the concept of model-driven DevSecOps, as introduced by Ponsard et al. (page 33), demonstrates the application of an internal model-based analysis and automation approach together with the external threat intelligence sharing for attack prevention, detection and recovery. Cobley et al. (page 34) pick up the specific topic of mobile forensics and introduce a standardised approach to forensics investigations, as well as accompanying training that is developed in the course of the European FORMOBILE project. Alexakos et al. (page 35) discuss a cybersecurity solution for a very special application context. They particularly focus on procedures for attack propagation monitoring and root cause analysis in Internet-of-Vehicle ecosystems.

Education and awareness

The last group of articles deals with educational initiatives and measures to raising awareness. Since criminals often target people instead of technology, e.g., in phishing campaigns, it cannot be stressed enough how important it is to educate people of cyber risks. Luh et al. (page 36) introduce PenQuest, a digital multi-player game that allows users to emulate cyberattacks on a game board. It is intended to assist risk assessment, support the reconstruction of adversarial events, and gamify security education. Hense et al. (page 38) introduce an approach to an educational Cyber Defence Centre (CDC) that trains students in a simulated environment where they gain skills in detecting attacks, closing vulnerabilities, and responding to security breaches in a realistic setting. Bassi et al. (39) focus specifically on schools and provide cybersecurity awareness-raising solutions, such as games and workshops, for pupils of all ages.

As the nature, motivation and targets of adversarial activities are quite diverse, we require a broad arsenal of counter measures. This issue of ERCIM News introduces many promising concepts, methodologies and solutions helping us to withstand and fight cybercrime activities. We must keep in mind that adversaries need to discover and successfully exploit only one vulnerability, being it technical or organisational, while defenders need to fix all of them to stay safe. Furthermore, with the introduction of novel computing paradigms, emergence of new technologies and IT's pervasion of almost all aspects of our lives, we must aim to improve current solutions and conduct further research to adapt them to a changing world.

Links:
[L1] https://kwz.me/hje
[L2] https://kwz.me/hjP
[L3] https://csirtsnetwork.eu/

Please contact:
Florian Skopik
AIT Austrian Institute of Technology
florian.skopik@ait.ac.at

Kyriakos Stefanidis
Industrial Systems Institute (ISI)
stefanidis@isi.gr

CC-DRIVER: Understanding the Technical Drivers of Cybercrime

by Evangelos Markatos (FORTH and University of Crete), Mary Aiken, Julia Davidson (University of East London), Alexey Kirichenko (F-Secure Corporation) and David Wright (Trilateral Research)

Over the past few years, we have seen cybercrime rising to become a trillion-dollar business worldwide. Although the cost of cybercrime was close to $5.5 trillion in 2020 [1], it is now estimated to double by 2025 [2]. To put this number in perspective, a cost of $10.5 trillion a year is $28 billion per day, or $20 million a minute, or close to $330,000 a second. At such staggering rates, it is imperative to understand the drivers of cybercrime and how they can be mitigated.

Our study on the technical drivers of cybercrime, within the EU-funded CC-DRIVER [L1] project, suggests that our digital society and attendant technical developments offer immense opportunities for cybercriminals. Take, for example, the proliferation of IoT devices that surround us: smart watches, smart phones, smart light bulbs, smart coffee makers, smart vehicles, smart homes, smart cities and, in fact, "smart everything". All of these smart entities have computing and communicating capabilities, which present a wide array of opportunities and attack vectors for cybercriminals. These smart devices are entry points to a private internal communication network in a home. If any of these smart devices are unprotected, cybercriminals can compromise the device and, from that one, they will be able to enter the home network. They may be able to probe the front door, scan the household PC, surveil the local traffic, deploy man-in-the-middle attacks, and create an account to permanently establish their presence – an insider threat in the home. These smart devices are "digital steppingstones" to a well-mounted attack. Notably, our affinity and increasing dependence on all things digital continue to increase the number of steppingstones.

As an example of technical drivers of cybercrime, we can consider the ever-evolving phenomenon of cryptocurrencies. Over the past few years, people have started to use cryptocurrencies not only to transfer money, but also as an alternative form of investment. Unfortunately, the anonymity, or indeed the pseudonymity, provided by the main crypto coins creates an ideal vehicle for cybercriminals to use crypto currencies for their financial transactions. In many aspects, cryptocurrencies are as good as, or sometimes even better than, cash. They are anonymous, they can be sent all over the world instantly, they do not involve any physical transfer of bills, and they leave few traces.

To protect their legal right to privacy online, legitimate network users use various mechanisms, such as anonymising networks and/or services that can provide some form of privacy. Such mechanisms include incognito browsers, cookie blockers, Virtual Private Networks (VPNs), and anonymising networks (such as Tor – 'the onion router' and entry point to the dark web). Although such systems are helpful for protecting the privacy of legitimate users, they can also be abused by cybercriminals to operate anonymously. This anonymity, especially as provided by Tor and similar networks, enables cybercriminals to hide their tracks and their illegal activities and, in doing so, evade detection by law enforcement authorities (LEAs).

In addition to the above and other technical drivers of cybercrime, our research has revealed a significant factor that has contributed to the current proliferation of cybercrime. Findings illustrate that the business model of cybercrime has moved from the "one-man-show", ad hoc type activity to organised "cybercrime-as-a-service", serious business operations. In the old days of cybercrime, one person, or a small number of closely collaborating people, were responsible for the entire cybercrime operation: from hacking into accounts all the way to shipping stolen goods. Things are different today. As cybercrime has become increasingly organised and tech-enabled, cybercriminals have started to specialise and become experts in niche areas. This has now evolved to a point where one cybercriminal might be an expert at hacking, another at recruiting, yet another at money-muling, and so forth. In this new operational scenario, everybody has become an expert at something, but few perpetrators are expert at everything. Cybercrime-as-a-service has, therefore, evolved as a means to help cybercrimi-

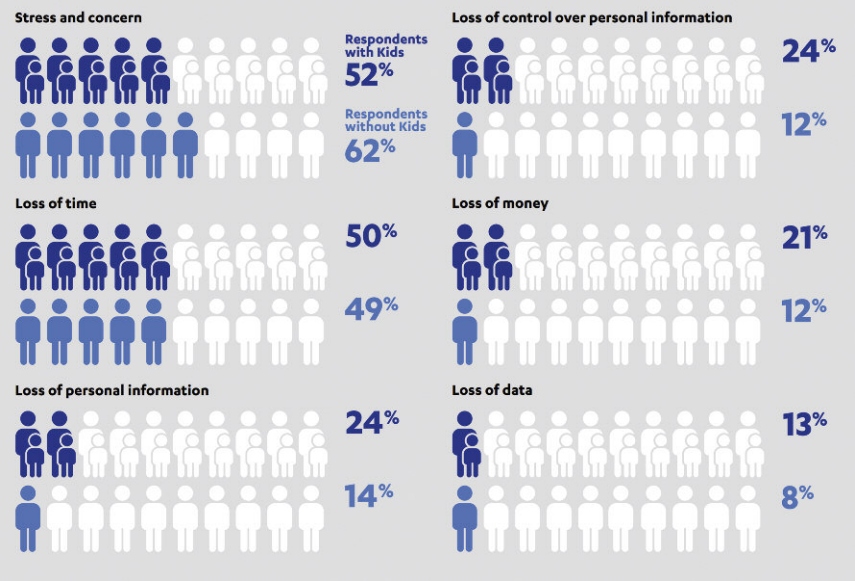


Figure 1: Effects of Cybercrime experienced by victims (source: F-Secure [L2]).

nals with different expertise collaborate in their criminal enterprise, combining and syndicating skillsets to create a whole operation that is greater than the sum of its parts. As soon as the cybercrime-as-a-service model started to gain traction and deliver results, increasing numbers of cybercriminals rushed to offer their services. Undoubtedly, this cybercrime model was facilitated by "offender convergence settings" in cyberspace, that is, dark nets in the dark web, for example: "hacking-as-a-service", offered to hack into accounts; "bulletproof-hosting-as-a-service", offered to host illegal operations; and "DDoS-attacks-as-a-service", offered to attack remote computers.

All of these technical drivers and the novel "as-a-service" business model make it evident that cybercrime is a highly adaptable, agile and growing business with new threat actors and criminal groups continuously entering the field. Understanding the models and the drivers is of the utmost importance for successfully countering cybercrime.

CC-DRIVER has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 883543. The views expressed in this article are those of the authors only and are in no way intended to reflect those of the European Commission.

Links:
[L1] https://www.ccdriver-h2020.com/
[L2] https://kwz.me/hfL

References:
[1] European Commission and High Representative of the Union for Foreign Affairs and Security Policy: "The EU's Cybersecurity Strategy for the Digital Decade", Joint Communication to the European Parliament and the Council, Brussels, 16 Dec 2020, p.3.
[2] S Morgan: "Cybercrime to Cost the World $10.5 Trillion Annually By 2025", Cybercrime Magazine, 13 Nov 2020. https://kwz.me/hf2

Please contact:
Evangelos Markatos, FORTH-ICS
markatos@ics.forth.gr

Countering Terrorist Financing

by Ross King (AIT Austrian Institute of Technology GmbH), Georgios Kioumourtzis (IANUS Consulting) and Georgios Th. Papadopoulos (FORTH-ICS)

The European Union's Internal Security Fund (ISF) will contribute to ensuring a high level of security in the Union, by supporting actions that help to prevent and combat terrorism and radicalisation, serious and organised crime, and cybercrime. One such project that has launched in January 2022 is Anti-FinTer: Versatile artificial intelligence investigative technologies for revealing online cross-border financing activities of terrorism.

Directive (EU) 2019/713 [1] on combating fraud and counterfeiting of non-cash means of payment points out that such means of payment are threats to security and enablers for other criminal activities, such as terrorism. The need for improving law enforcement capacity and developing expertise in the area of terrorist financing is also in line with the Financial Action Task Force (FATF) Report on Terrorist Financing [L1], which calls for deepening the understanding of financing mechanisms. The need for action in Europe is also clear, for example through the terrorist modi operandi identified in Europol's recent IOCTA [2] and SOCTA [L2] reports: First, terrorists are implementing crowdfunding campaigns to collect resources to finance their activities. In an attempt to maintain anonymity, terrorist crowdsource platforms combine cryptocurrency and Dark Web market technologies. Second, terrorist groups aiming at trafficking drugs and/or firearms in large quantities tend to adopt a hierarchical internal structure that is typically based in multiple countries inside and outside the borders of the EU. Illegal transfers and smuggling techniques are typically implemented with and covered by conventional legal business activities. Third, an emerging scenario has been detected in which terrorists aim to collect revenues from selling extremist versions of common products (e.g., merchandise like t-shirts or flags that market or promote extremist groups) or other illicit goods (e.g., counterfeits, illegal drugs, or weapons) to the general public or other extremists/terrorists. These activities are often realised on the surface web but may also involve links to Dark Web platforms or markets.

The new ISF project Anti-FinTer [L3] will improve law enforcement capabilities, increase capacity, and develop expertise in the area of terrorist financing associated with activities in the Dark Web, crypto-assets, new payment systems and darknet marketplaces. The project consortium consists of ten European partners: four research organisations, two small-to-medium enterprises, and four law enforcement agencies (LEAs). The project has launched in January 2022 and will run for two years.

Three distinct project activities will contribute to the combat against terrorism and cybercrime. The first is through the facilitation of knowledge exchange among stakeholders and the documentation of best practices, and through risk analysis and policy recommendations in workshops and virtual meetings. The second is through the integration of existing forensic software to create a Toolkit for training investigators and analysts in new investigative techniques that include crypto-asset analysis, new payment channels such as the Lightning Network, text and image analysis from surface web, dark web and social media channels to identify common actors and correlate terrorist activity with cryp-

tocurrency transactions, and artificial intelligence analytics for detecting transaction anomalies. The third is through the development of training curricula and an exercise environment used in virtual and face-to-face training events that will be organised and carried out during the project along with train-the-trainer events that will ensure a wider impact for the curricula.

Financial investigations typically include multiple, iterative, and refined data collection and analysis steps, which involve financially related information (e.g., transactions, purchase records), but also associated context (e.g., social media analysis, terrorist propaganda incidents, ransomware and phishing reports). Anti-FinTer will extend the capabilities of investigations to include the context of transactions in Dark Web markets, as explored in the H2020 ASGARD [L4] project.

A well-established methodology that has been validated in various financially related crimes is the so-called "follow the money" approach, where the fundamental principle states that tracking the flow of money will very likely lead to the detection and identification of the suspects behind the illicit activities. Anti-FinTer extends the capabilities of financial investigations to include money flows in the form of crypto-assets and new payment systems, by building on results from the H2020 TITANIUM [L5] project.

However, the "follow the money" approach alone cannot guarantee the successful completion of an investigation. It has been observed that the organised criminal groups tend to fragment their business activities, in an attempt to obfuscate LEA operations. Therefore, a "follow the actor" approach, as illustrated in Figure 1, will also be developed during the course of the project, aiming at identifying the (groups of) actors behind different types of crime and various related activities (e.g., firearms trafficking), combining both cyber and physical information cues. This methodology combines information about the location and identity of suspects with virtual information from open-source intelligence, such as cryptocurrency ledgers, public and private sector archives, and the Internet. The approach puts the focus on jointly analysing multiple illegal financial investigations/cases, to reveal the identities of the same (group of) actors that are behind all these incidents. For example, the application of visual analytics applied to dark web data in the context of Anti-FinTer will enable LEAs to pursue more effectively the "follow the actor" approach by identifying commonalities in the depiction of illicit goods.

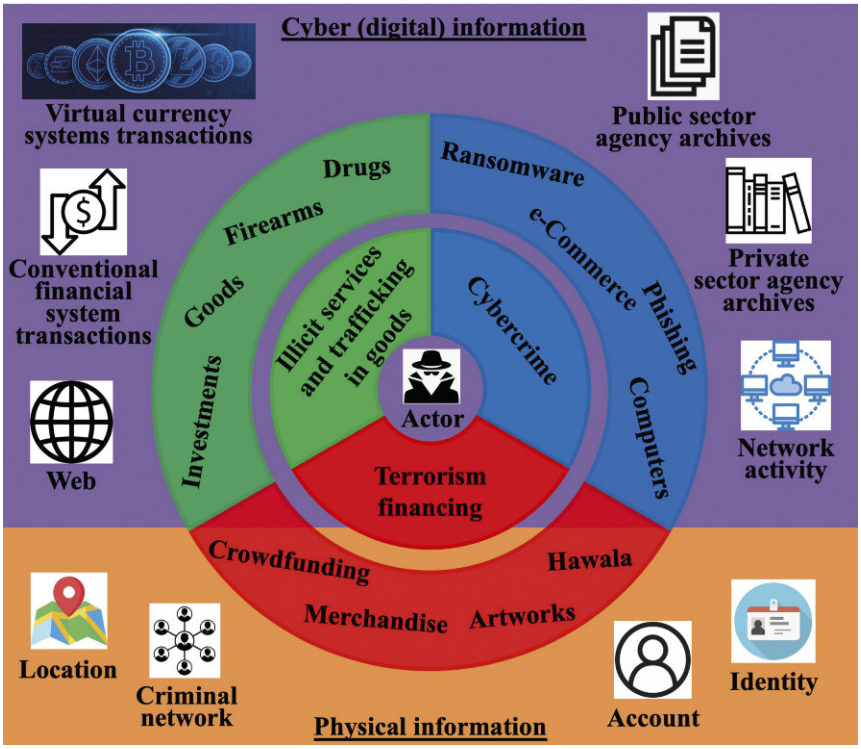


Figure 1: Anti-FinTer's proposed "follow-the-actor" financial investigation strategy.

This report was funded by the European Union's Internal Security Fund — Police under Grant Agreement No. 101036262. The content of the report represents the views of the authors only and is their sole responsibility. The European Commission does not accept any responsibility for use that may be made of the information it contains.

Links:

[L1] https://kwz.me/hfJ
[L2] https://kwz.me/hfM
[L3] https://anti-finter.eu/
[L3] https://www.asgard-project.eu/
[L4] https://www.titanium-project.eu/

References:

[1] Directive (EU) 2019/713 of the European Parliament and of the Council of 17 April 2019 on combating fraud and counterfeiting of non-cash means of payment and replacing Council Framework Decision 2001/413/JHA. https://kwz.me/hfW
[2] Europol (2021), Internet Organised Crime Threat Assessment (IOCTA) 2021, Publications Office of the European Union, Luxembourg. https://kwz.me/hfQ

Please contact:

Ross King
AIT Austrian Institute of Technology GmbH, Austria
ross.king@ait.ac.at
Anti-FinTer Coordinator

Georgios Kioumourtzis
IANUS Consulting, Cyprus
gk@ianus-consulting.com
Anti-FinTer Dissemination Manager

Georgios Th. Papadopoulos
Foundation for Research and Technology – Hellas, Greece
gepapado@ics.forth.gr
Anti-FinTer Technical Manager

Mitigating Financial Cybercrime with BPMN-based Standardised Investigation Procedures

by George Tsakalidis, (Financial and Economic Crime Unit - S.D.O.E. (Operational Directorate of Macedonia)) and Kostas Vergidis, (University of Macedonia)

A systematised investigation process for copyright-related cybercrime offences has been designed in the Business Process Model and Notation (BPMN) and implemented by financial crime investigators of a Law Enforcement Agency. The proposed approach has increased the efficiency of the performed investigations and the dissemination of knowledge to relevant agencies.

The advances in information systems and network technologies have created a novel conducive environment for criminal activity, broadly known as cybercrime. An emerging cybercrime type with high frequency of occurrence and severe consequences for the global economy, refers to offences related to infringements of copyright and related rights [1]. In the financial context, this critical cybercrime covers the illegitimate download, copy, distribution, and usage of intellectual property, also in the form of software, programs, and tools. In Greece, this cybercrime is extensively committed by individuals, companies, and organisations, recording an average of 61% of unlicensed software installations during the period 2011–2017 [2]. The presented project is an initiative for the standardisation of the investigation procedure of this widespread cybercrime and the subsequent improvement of the efficiency of a Law Enforcement Agency. The investigation process refers to the National Law N2121/1993 and the Intellectual Property (IP) infringements committed through the usage of illegitimate (pirate) software products by companies and organisations in Greece.

The institutions involved in the research project are: (a) the Financial and Economic Crime Unit - S.D.O.E. (Operational Directorate of Macedonia) of the Greek Ministry of Finance and (b) the Department of Applied Informatics of the University of Macedonia, Thessaloniki, Greece. The investigators of the Agency were involved in recording and specifying the investigation activities and the different scenarios that can emerge, interrelating the applying laws and checking process conformance. Researchers from the University of Macedonia were assigned with modelling the investigation process using a state-of-the-art standard, the Business Process Model and Notation (BPMNv2.0) [3].

The established BPMN notation was adopted to be readily understandable by all users, from the analysts that create the initial process drafts, to the developers that implement the technology, and finally, to the investigators that manage and execute these processes. The steps followed during the standardisation of the investigation procedure involved the cooperation of the two institutions, to translate fragmentary investigation steps into an executable BPMN diagram. The process (Figure 1) initiates with the arrival of the investigators to the company and the display of their IDs and investigation mandate. During the personal computer (PC) software investigation subprocess, the investigators search the company and set all PCs in operation. For unveiling the installed software, both portable auditing software and manual search methods are used. The installed software programs are documented, and the list is provided to the company for collecting both the evidence of legitimate acquisition and usage.

Depending on whether the presented evidence refers to the installed software, the unit will either complete the investigation, or proceed with the enforcement of sanctions determined in N.2121/1993 (followed by composition of official documents and software uninstall). If the illegal software programs number more than 50, the investigators proceed to arrest the company owner or representative, initiate the penal sanction procedure and forward a copy of the offence ascertainment document to the local revenue office (AADE) for collecting the imposed fine. In the case of fewer than 50 programs, the unit notifies the owner or representative of the capability to cease penal sanction in case the fine is paid within the next 24 hours. For fewer than 50 software programs, the penal sanction is ceased in the case of punctual payment and initiated in the case of non or overdue payment. The final steps involve forwarding the investigation results to the Hellenic Copyright Organisation (OPI) for notifying the software manufacturers and to the AADE office for further investigation.

The use of the investigation process provides a multitude of benefits for both the agency and investigators. In particular, this is a typical example of digital transformation in the public sector since it fulfils: (a) the need for compliance with the Digital Transformation Strategy 2020-2025 of Greece, (b) the obligation of public agencies to index their administrative processes to the National Process Registry (Law N.4727/2020), and most importantly (c) the need to standardise the investigation processes of the agency and to enhance the efficiency and situational awareness of all the investigators that are tasked with this authority.

Regarding the latter, the investigators are accustomed to the different scenarios that can emerge, they know beforehand which official documents to complete, and they are aware of the applying laws and rules. By using the process, investigators examine more cases in the same time periods, a fact that has reduced the operating expenses of the agency, both monetary and in human resources. The orientation of the project is the application of the systematised investigation process by the rest of the investigators of the Directorate. The process has been distributed to the rest of the investigators to get accustomed to the methodology and apply the process in everyday cases. A principal advantage of the process is that it can be readily applied by investigators

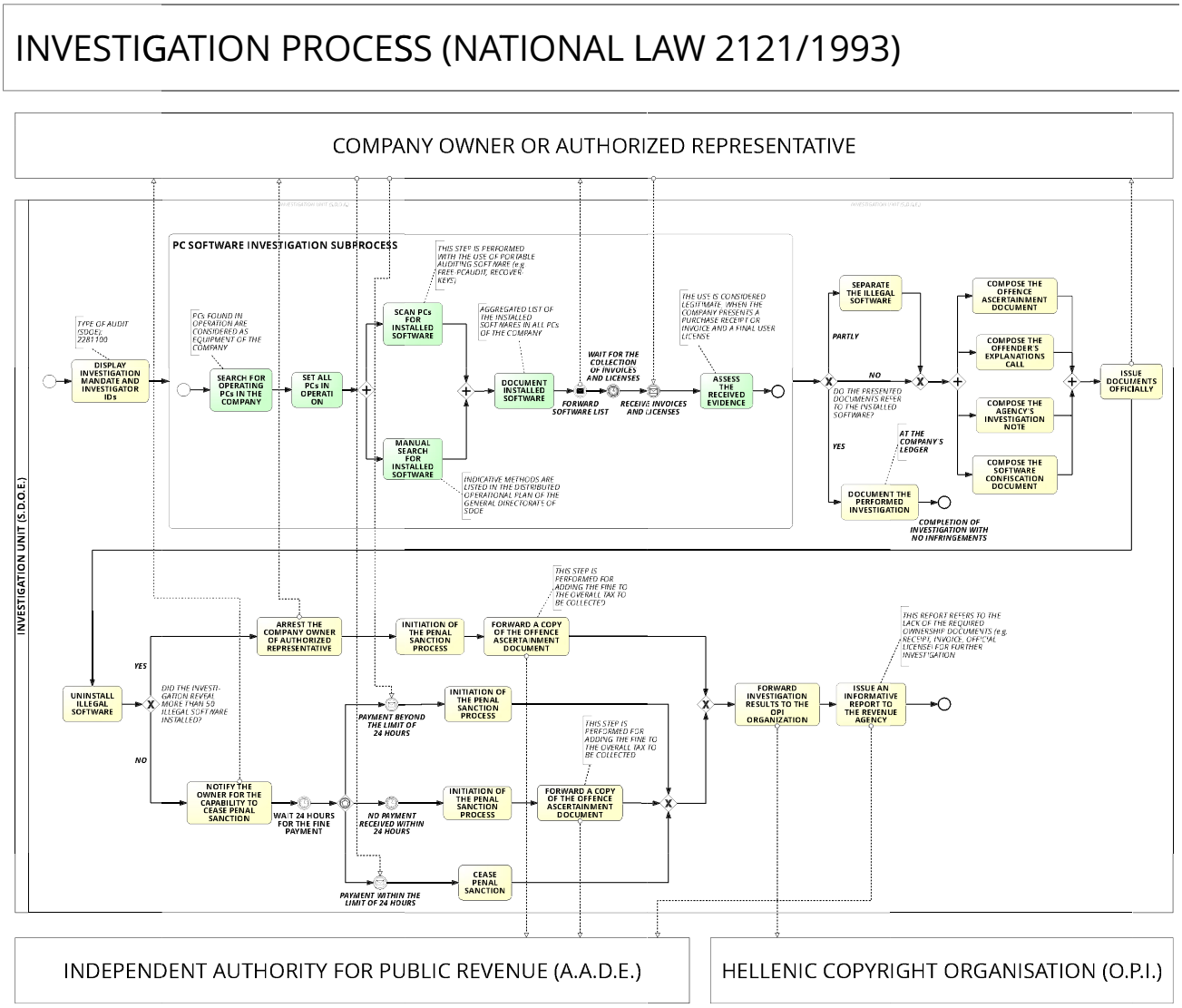


Figure 1: Investigation Process for IP infringements related to illegal software usage (National Law N.2121/1993).

of the Operational Directorate S.D.O.E. of Attica and other Law Enforcement Agencies that have the same jurisdiction.

There is ongoing work for further extending the project, primarily in quantifying the efficiency of the investigation process. The authors are creating an inventory with statistical metadata regarding execution times and number of investigators involved, in an attempt to identify defective process parts and bottlenecks through simulation. In this manner, the authors intend to redesign the process and optimise critical performance criteria, like execution time and/or cost, process flexibility and overall external quality.

References:

[1] G. Tsakalidis and K. Vergidis: "A systematic approach toward description and classification of cybercrime incidents", IEEE Trans. Syst. Man Cybern. Syst., vol. 49, no. 4, pp. 710–729, 2017.

[2] BSA, The Software Alliance: "BSA Global Software Survey," Washington, DC, Annual, Jun. 2018. Accessed: Feb. 13, 2022. [Online]: https://kwz.me/hfF

[3] Object Management Group (OMG): "About the Business Process Model And Notation Specification Version 2.0," Dec. 2013. Accessed: Jan. 23, 2022. [Online]. Available: https://kwz.me/hfG

Please contact:

George Tsakalidis, Financial Crime Investigator, Financial and Economic Crime Unit - S.D.O.E. (Operational Directorate of Macedonia), Greece
giorgos.tsakalidis@uom.edu.gr

Digital Forensics for the Detection of Deepfake Image Manipulations

by Sara Ferreira (University of Porto), Mário Antunes (Polytechnic of Leiria) and Manuel E. Correia (University of Porto)

Tampered multimedia content is increasingly being used in a broad range of cybercrime activities. The spread of fake news, misinformation, digital kidnapping, and ransomware-related crimes are among the most recurrent crimes in which manipulated digital photos are being used as an attacking vector. One of the linchpins of accurately detecting manipulated multimedia content is the use of machine learning and deep learning algorithms. This work proposed a dataset of photos and videos suitable for digital forensics, which has been used to benchmark Support Vector Machines (SVM) and Convolution Neural Networks algorithms (CNN). An SVM-based module for the Autopsy digital forensics open-source application has also been developed. This was evaluated as a very capable and useful forensic tool, winning second place on the OSDFCon international Autopsy modules competition.

Cybercrime is challenging national security systems all over the world, with malicious actors taking advantage of and exploiting human and technical vulnerabilities. The widespread global reach of cyberattacks, their level of sophistication and impact on society have also been reinforced by the pandemic we are all currently enduring. This has raised a global awareness of how dependent we now are on the Internet to carry out normal daily activities, and how vulnerable we all are to fraud and other criminal activities in cyberspace.

Deepfakes use artificial intelligence to replace the likeness of one person with another in video and other digital media. They can inflict severe reputational and other kinds of damage to their victims. Coupled with the reach and speed of social media, convincing deepfakes can quickly reach millions of people, negatively impacting society in general. Deepfake attacks may have different motivations like fake news, revenge porn, and digital kidnapping, usually involving underage or otherwise vulnerable victims and possibly associated with ransomware blackmailing. Digital forensics analysis of such cases, when conducted manually and solely by the means of a human operator, can be very time-consuming and highly inefficient in identifying and collecting complete and meaningful digital evidence of cybercrimes, often because of the misclassification of files.

Effective forensic tools are essential, as they have the ability to reconstruct evidence left by cybercriminals when they perpetrate a cyberattack. However, an increasing number of highly sophisticated tools make life much easier for cybercriminals to carry out complex digital attacks. The criminal investigator is thus faced with a very difficult challenge in trying to keep up with these cyber-criminal operational advantages. Autopsy [L1] () is a digital forensics tool that helps to level the field. It is open-source and widely used by criminal investigators to analyse and identify digital evidence of suspicious activities. Autopsy incorporates a wide range of native modules to process digital objects, including images (on raw disks), and it allows the community to develop more modules for more specialised forensic tasks.

Machine Learning (ML) has boosted the automated detection and classification of digital artifacts for forensic investigative tools. Existing ML techniques to detect manipulated photos and videos are seldom fully integrated into forensic applications. Therefore, ML-

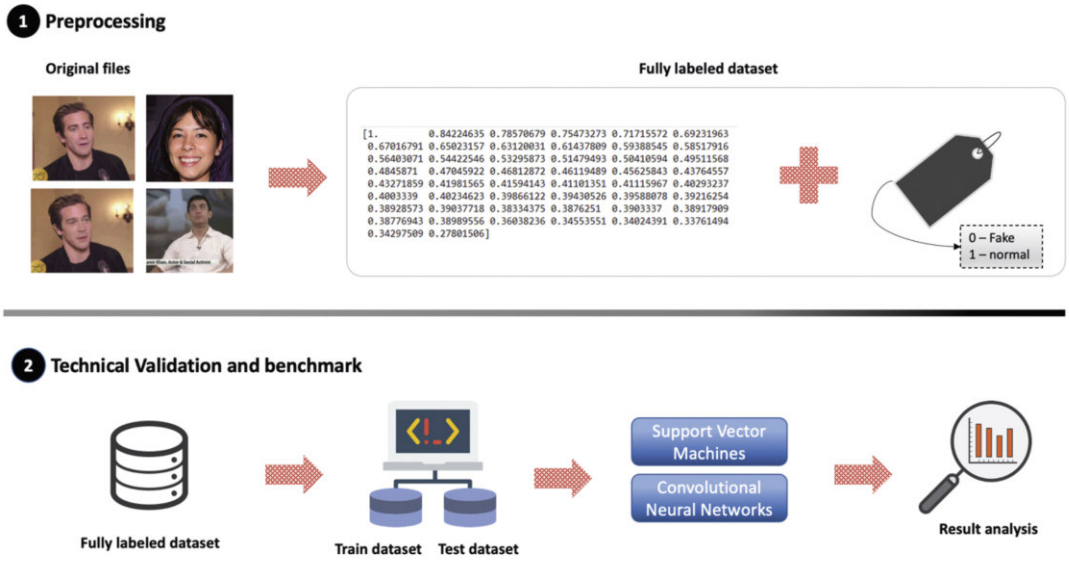


Figure 1: Overall architecture of the preprocessing and technical validation of the dataset.

based Autopsy modules, capable of detecting deepfakes, are relevant and will most certainly be very much appreciated by the investigative authorities [1]. Well proven ML methods for deepfake detection have not yet been fully translated into substantial gains for cybercrime investigation, as those methods have not often been incorporated into the most popular state-of-the-art digital forensics tools. In this work [1, 2, 3] we made the following contributions:

- A labelled and balanced dataset composed of about 52,000 examples of genuine and manipulated photos and videos that incorporates the most common manipulation techniques, namely splicing, copy-move and deepfaking [3].
- An SVM-based model capable of processing multimedia files and detecting those that were digitally manipulated. The model processes a set of simple features extracted by applying a Discrete Fourier Transform (DFT) method to the input file.
- The development of two ready-to-use Autopsy modules to detect the fakeness level of digital photos and input video files, respectively [L2, L3].

The overall architecture is shown in Figure 1. It is composed of two main stages: pre-processing, and technical validation and benchmark of the dataset [2].

Pre-processing consists of reading the photos and taking up to four frames per second from the input videos through the OpenCV library. By having all the photos in the dataset, the features' extraction is made by applying the DFT method to generate labelled input datasets for both training and testing. The processing phase corresponds to the SVM and CNN processing. The implementation of SVM processing was made through the scikit-learn library for Python 3.9. The model created by SVM at the processing phase, is used to get a "fake" score for each photo in the testing dataset. The tests were carried out with a 5-fold cross-validation, by splitting the dataset into five equal parts and using four for training and one for testing. These two phases were incorporated in a developed standalone application, which was further integrated as two separated Autopsy modules [L4].

The deliverables obtained with this research, namely the ready-to-use Autopsy modules, can give a helping hand to digital forensics investigators and leverage the use of ML techniques to fight cybercrime activities that involve multimedia files. The low processing time and the high performance obtained with the DFT-SVM method make it eligible to be incorporated as a plugin that may be used easily and in real time, to detect the fakeness level of multimedia content spread in social networks.

Links:

[L1] https://www.autopsy.com/
[L2] https://kwz.me/hfl
[L3] https://kwz.me/hfp
[L4] https://kwz.me/hfg

References:

[1] S. Ferreira, M. Antunes, M.E. Correia: "Forensic Analysis of Tampered Digital Photos", in: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2021, Springer LNCS 2021, vol 12702, pp. 461-470. https://doi.org/10.1007/978-3-030-93420-0_43

[2] S. Ferreira, M. Antunes, M.E. Correia: "Exposing Manipulated Photos and Videos in Digital Forensics Analysis", Journal of Imaging, 2021; 7(7):102. https://doi.org/10.3390/jimaging7070102

[3] S. Ferreira, M. Antunes, M.E. Correia: "A Dataset of Photos and Videos for Digital Forensics Analysis Using Machine Learning Processing", Data, 2021; 6(8):87. https://doi.org/10.3390/data6080087

Please contact:

Manuel E. Correia
University of Porto, Portugal
mdcorrei@fc.up.pt

Privacy-Preserving Collaborative Anomaly Detection to Fight Cybercrime

by Rudolf Mayer (SBA Research)

Anomaly detection is an important part of countering cybercrime, by detecting e.g., fraud or intrusions. Especially with an ever-growing amount of data (such as logs or transactions) being collected, automated analysis of these data for malicious behaviour becomes essential. In several settings, such analysis might be performed by third parties or be collaborative, to learn from more and diverse experiences by different collaborators. Thus, means to access such often confidential data in a privacy-preserving manner are required. Collaborative Learning and synthetic data are two promising approaches to fulfil this purpose.

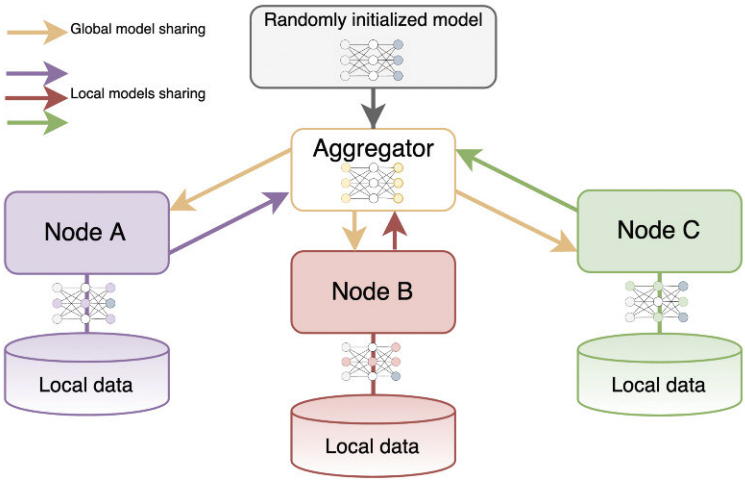
The demand for and practice of data sharing and exchange between different data collecting parties is increasing, often because different data sets complement each other, or because the processing and analysis of data is outsourced. Many interesting knowledge discovery tasks are dependent on large, high-quality amounts of data, and anomaly detection, e.g., for the purpose of detecting cybercrime, is no exception – especially as fraudulent behaviour changes over time.

However, when data is sensitive, e.g., when it concerns individuals or is business related, there are certain regulatory and other barriers for data sharing and collaboration. Still, collaborative analysis of data can be very beneficial, as different organisations might be affected (targeted) at different stages. Thus, learning from misuse patterns that other parties have already been exposed to, such as network intrusions, financial fraud, malware, or other forms of cybercrime, can be extremely valu-

able. Therefore, means to enable such data exchange collaboration are required. Anonymisation of data is one frequently studied approach, e.g., in the form of k-anonymity or differential privacy. However, k-anonymity has been shown to be still prone to linkage attacks when adversaries have background knowledge and access to other data sources, and differential privacy is not easily applicable to all types of analysis techniques and methods. Therefore, alternative approaches such as synthetic data generation and federated learning have also been explored, and recently specifically evaluated for anomaly detection.

Synthetic data is generally considered as data obtained not from direct measurement. In the context of data analysis efforts, it is often considered to be data generated (or synthesised) from a real dataset that cannot be shared e.g., for privacy considerations. The aim of synthetic data in this context is to generate a dataset containing records that are similar to the original ones, and thus allow a similar analysis as the original data, without actually disclosing real, single data points – which might be data points that contain sensitive information that is not to be disclosed. Synthetic data has successfully been shown to achieve good results, e.g., for classification and regression tasks. However, one could assume that synthetic data might not be easily usable for a task such as anomaly detection, which deals with outliers, while synthetic data generally preserves global characteristics. It thus might tend to represent rather the ones of the legitimate, normal cases, and might fail to generate representative anomalies.

However, in a recent work [1], we evaluated three different approaches to generate synthetic data for supervised, semi-supervised and unsupervised anomaly detection settings, including credit card fraud. While anomaly detection is a hard task especially for the latter two, synthetic data reaches similar effectiveness as the models trained on the original data, and can thus be considered a viable alternative: instead of sharing and centralising real data, different collaborators can exchange synthetic data generated locally.
Another approach for privacy-preserving data analysis is federated learning [2], which is especially useful



Federated learning with three nodes learning models on their local data, and model averaging by a central aggregator.

in settings where data is collected in several distributed locations. In federated learning, first models are trained locally on each data source, before they are aggregated to a common, global model; this is usually repeated for a few cycles, to allow convergence. Thus, the training data remains at the source, and the only type of information exchanged are the model parameters – which generally represent a strong abstraction of the local model. Aggregation strategies vary for each anomaly detection method. For example, for the supervised task with neural networks, e.g., simple averaging of the learned model weights (or from the gradients to adapt those) is a suitable strategy, while for other methods, different aggregation strategies are required, or ensembles could be built. A recent evaluation shows that especially supervised, but also semi-supervised methods can achieve comparable detection rates [3]. In general, federated learning is heavily influenced by the way data is distributed among the clients, and model aggregation needs to consider if there is an imbalance and skewness in the amount of data and anomalies present at each local site, and more advanced aggregation strategies might need to be developed for these cases.

Protecting the confidentiality of training data is an important aspect in many settings, but recent approaches have shown promising results in several anomaly detection settings, including in cybercrime. While it might not be always possible to replicate results that would be achievable in an idealised approach of centralising all data, collaboration can improve the results clients would achieve if just leveraging their own data.

References:

[1] R. Mayer, M. Hittmeir, A. Ekelhart: "Privacy-preserving anomaly detection using synthetic data", in Proc of the 34th Annual IFIP WG 11.3 Conf. on Data and Applications Security and Privacy (DBSec), Regensburg, Germany, 2020. Springer.

[2] P. Kairouz, H. Brendan McMahan, et al.: "Advances and Open Problems in Federated Learning", Foundations and Trends in Machine Learning, 14(1-2), 2021.

[3] F. Cavallin, R. Mayer: "Anomaly Detection from Distributed Data Sources via Federated Learning". Proceedings of the 36th International Conference on Advanced Information Networking and Applications (AINA), Sydney, Australia, 2022. Springer International Publishing.

Please contact:

Rudolf Mayer
SBA Research, Austria
rmayer@sba-research.org

Tiny Machine Learning: A New Technique for AI Security

by Hui Han (Fraunhofer IESE) and Jingyue Li (NTNU)

Tiny machine learning (TinyML) is the intersection of machine learning (ML) algorithms and embedded systems (hardware and software) in terms of low latency, low power, and small size. It allows data to be kept mainly on edge devices and to be processed and have ML tasks run directly in the device. Therefore, the TinyML paradigm is expected to preserve AI security and combat cybercrimes. In this study, we explore how TinyML, the cutting-edge of ML technologies, solves relevant AI security problems (including cybercrimes) from the AI lifecycle aspect: data engineering, model engineering and model deployment. Finally, we discuss the opportunities for future research.

Tiny machine learning (TinyML) is a fast-growing field of machine learning (ML) technologies and applications that include algorithms, hardware (dedicated integrated circuits), and software that can perform on-device sensors (vision, audio, inertial measurement unit, biomedical, etc.) data analytics at extremely low power, typically in the order of milliwatts, enabling a variety of always-on machine learning use cases on battery-powered devices [L1]. TinyML will play an essential role in our daily interactions with ML in the near future.

TinyML for AI security

The TinyML system is the integration of ML-based mechanisms with Microcontroller Units (MCUs) based edge devices. This smooths the path for the development of efficient services and novel applications that do not need ubiquitous processing support from the cloud [1]. TinyML can make the data stay on-premise, which enhances security and ensures data privacy without letting sensitive raw data leave devices. More importantly, TinyML can enable data analysis and real-time decision-making in the field without relying on the computing power from a cloud, which preserves AI security.

We explore how TinyML as a new technique solves relevant AI security problems from the AI lifecycle aspect: data engineering, model engineering and model deployment.

Data engineering

Data privacy and security in the digital age is a significant issue in AI security. Transmitting raw data through unstable and lossy wireless channels from edge devices to the cloud can jeopardise data privacy or lead to stolen data, data loss or compromised data, transmission errors and cyberattacks (e.g., man-in-the-middle (MITM) attack) [1]. TinyML allows embedded devices to process data locally (close to the sensor), which results in better data privacy and security.

Model engineering

The security challenges for model engineering are very important. Running ML models (deep-learning models or neural networks models) with TensorFlow Lite on ultra-low-power microcontrollers – e.g., SparkFun Edge Development Board Apollo3 Blue and Arduino Nano 33 BLE Sense board – for ML inferences, TinyML offers multiple advantages (low cost and energy and ubiquitous MCUs for ML models), notably preserving AI security.

Model deployment

The last stage of the lifecycle refers to the deployment of the model in practical use. TinyML makes responding or making decisions in a short time possible by avoiding continuous connectivity to the cloud. This brings a wide array of use cases (such as device identification, authentication, intrusion detection, malware detection, anomaly detection, secure range search, and attack detection) of TinyML deployment where privacy and security are considered vital effect factors.

Cybercrimes such as network intrusion, malware and human intervention are quite widespread in this digital society. As mentioned above, the significant applications of TinyML are intrusion detection, malware detection and attack detection, which can combat some cybercrimes.

Future work

TinyML is at the cutting edge of computer technology and AI, meaning it faces many challenges as well as more opportunities. We recommend some potential future research directions:

- Creating an effective TinyML dataset or repurposing existing datasets for TinyML.
  Traditional ML models need a large amount of data that are hard to get. Collecting and labelling these large datasets is expensive and the data may be only used for special tasks. Although there are a number of well-known open-source datasets for training ML models, these public datasets are not suitable to train ultra-low-power models for embedded devices, and are relatively large for TinyML-specific use cases. Therefore, a specific TinyML dataset is one of the promising areas for future research. TinyML scholars could set up brand new datasets or take advantage of these existing datasets by repurposing them.
- Designing specific algorithms for TinyML-specific requirements relevant to AI security.
  Most ML algorithms are very vulnerable to perturbations and TinyML algorithms. What's more, some current algorithms are either computationally too intensive or overly complex for TinyML deployment [2]. Therefore, one interesting research direction is to empirically evaluate how robust the existing TinyML models/algorithms are and how to improve their robustness in terms of AI security.
- Protocol, benchmarks, standards and rules for TinyML referring to AI security.
  New endpoint security mechanisms are not only required to meet adequate security standards, but also such mechanisms must be as lightweight as possible. The TinyML community extends the existing MLPerf benchmark suite to TinyMLPerf with TinyML systems. The goal of TinyMLPerf is to provide a detailed description of the motivation and guiding principles for benchmarking of TinyML systems [L2]. Although the security protocol Object Security for Constrained RESTful Environments (OSCORE) is designed to tackle this challenge, more standards are needed for AI security when employing TinyML [3].

The work described in this article has been carried out in the frame of an ERCIM "Alain Bensoussan" Fellowship.

Links:
[L1] https://www.tinyml.org
[L2] https://mlperf.org

References:
[1] R. Sanchez-Iborra and A. F. Skarmeta: "TinyML-enabled frugal smart objects: challenges and opportunities," IEEE Circuits and Systems Magazine, vol. 20, no. 3, pp. 4–18, 2020, doi: 10.1109/MCAS.2020.3005467.
[2] S. Siddiqui, C. Kyrkou, and T. Theocharides: "Mini-NAS: a neural architecture search framework for small scale image classification applications", in TinyML Research Symposium, 2021, pp. 1–8.
[3] H. Doyu, R. Morabito, and M. Brachmann: "A tinyMLaaS ecosystem for machine learning in IoT: overview and research challenges," in International Symposium on VLSI Design, Automation and Test, 2021, pp. 1–6. doi: 10.1109/VLSI-DAT52063.2021.9427352.

Please contact:
Hui Han
Fraunhofer Institute for Experimental Software Engineering IESE, Germany
hui.han@alumnos.upm.es

IoT Malware Detection with Machine Learning

by Levente Buttyán (Budapest University of Technology and Economics) and Rudolf Ferenc (University of Szeged)

Embedded devices are increasingly connected to the Internet to provide new and innovative applications in many domains. However, these IoT devices can also contain security vulnerabilities, which allow attackers to compromise them using malware. We report on our recent work on using machine learning for efficient and effective malware detection on resource-constrained IoT devices.

Embedded devices connected to the Internet are threatened by malicious programs (viruses, worms, also known as malware). One of the most infamous examples for IoT malware is Mirai, which infected hundreds of thousands of IoT devices and launched the largest distributed denial-of-service attack against Internet-based services in 2016, but the IoT threat landscape includes many other malware families as well.

Anti-virus products developed for traditional IT systems have higher resource needs than that offered by embedded IoT devices. The required amount of free storage space and memory to run these products is often measured in gigabytes, which exceeds the capacity of typical IoT devices, such as WiFi routers, IP cameras, smart household appliances, and wearable devices. In addition, many existing anti-virus products do not even support the operating systems used on IoT devices. Therefore, they could not be installed, even if a particular IoT device met their system requirements.

Since malware detection is essentially a classification task, machine learning-based methods have been applied in this area in recent years [1]. Machine learning-based malware detection has several advantages. For example, these methods are not only able to detect previously seen malware, but they can also detect new, previously unseen malware if it is similar in some way to previously seen samples. Another advantage is that machine learning models can represent more concisely the characteristics of previously seen malware patterns than the signature databases used in tradi-

tional signature-based detection. This makes machine learning-based malware detection particularly well-suited for resource-constrained environments such as embedded IoT devices.

Hence, in our projects (MILAB [L1] and SETIT [L2]), we work on new machine learning-based malware detection methods tailored for resource-constrained IoT devices. In a recent paper [2], we have proposed SIMBIoTA (SIMilarity Based IoT Antivirus), an effective and efficient anti-virus solution for such devices. The operating principles of SIMBIoTA are similar to those of traditional signature-based anti-virus solutions, but SIMBIoTA uses TLSH hash values of known malware instead of raw binary signatures for detection purposes. TLSH is a similarity hash algorithm, and it is different from cryptographic hashes: similar inputs result in similar TLSH hash values, and SIMBIoTA takes advantage of this feature. More specifically, embedded IoT devices that use SIMBIoTA store only a few TLSH hash values of known malware, and they compare the TLSH hash value of new files to these stored hashes. If the TLSH hash of an unknown file is similar to that of a known malware, the unknown file is detected as malware.

We evaluated the detection performance of SIMBIoTA and measured a true positive detection rate of more than 90% on average, even for previously unseen malware samples. Moreover, in the experiments performed, its false positive detection rate was 0%. In terms of resource needs, SIMBIoTA requires just a few tens of kilobytes of storage space, which is certainly available even on resource-constrained IoT devices.

In a follow-up work, we also used TLSH hash values for malware detection on IoT devices, but in a manner different from that of SIMBIoTA. Our key observation is that, thanks to their well-defined structure, TLSH hash values can be used as feature vectors for training machine learning models, which can then be used for malware detection. We call the resulting antivirus solution SIMBIoTA-ML. We showed that this approach can result in interesting trade-offs in terms of detection performance and resource usage on IoT devices. More specifically, SIMBIoTA has lower storage requirements and false positive detection rate than SIMBIoTA-ML, but SIMBIoTA-ML outperforms SIMBIoTA in terms of true positive detection rate, achieving more than 95% on average (see Figure 1). We also showed that SIMBIoTA's database of TLSH hash values increases over time, which has an impact on its detection time. Specifically, the larger the database is, the longer it takes for SIMBIoTA to decide whether an unknown file is malicious or not. By contrast, we showed that SIMBIoTA-ML has a near-constant running time, which allows for better estimation of the delay introduced by the anti-virus solution, and this can be an advantage in case of real-time applications (e.g., cyber-physical systems). In addition, using our DeepWater [3][L3] machine learning framework, we compared the detection

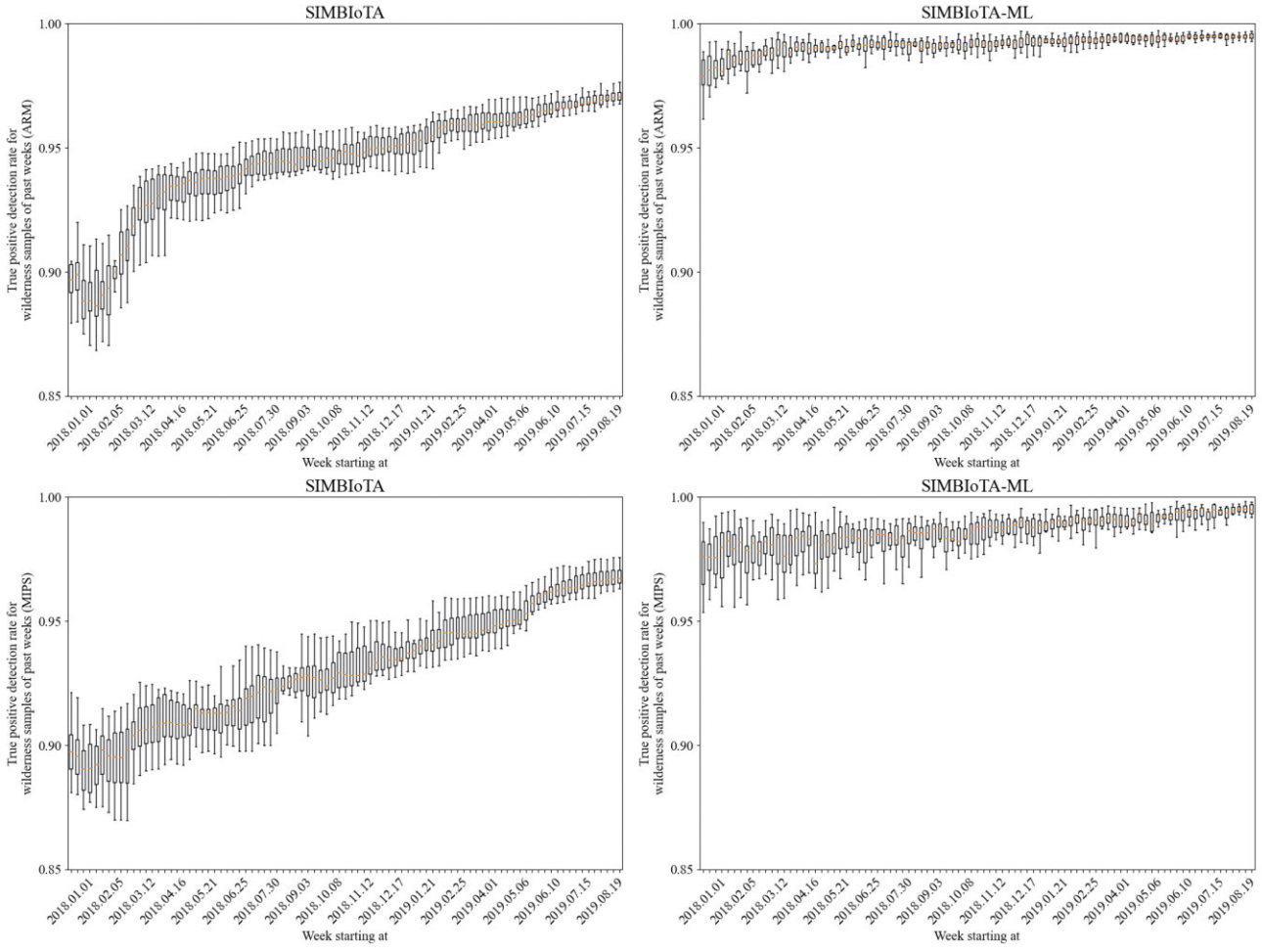


Figure 1: Box plot of the true positive detection rate of SIMBIoTA and SIMBIoTA-ML on ARM and MIPS samples.

performance of SIMBIoTA-ML when used with different machine learning models, and found that the best performance is achieved by the logistic regression model, which also turns out to be the least resource demanding in terms of memory usage and prediction time.

We performed all experiments using our IoT malware benchmark dataset called CUBE-MALIoT, which we made public at [L4]. This data set consists of 29,209 malicious samples developed for the ARM platform and 18,715 malicious samples developed for the MIPS platform. To the best of our knowledge, such a large dataset containing raw binaries of IoT malware was not previously available publicly to the research community. We hope that CUBE-MALIoT will become a de facto benchmark dataset in IoT malware detection in order to satisfy the need for the comparability and reproducibility of results of different research groups.

Links:

[L1] https://mi.nemzetilabor.hu/about-us
[L2] https://kwz.me/hfY
[L3] https://kwz.me/hfB
[L4] https://kwz.me/hfb

References:

[1] Ucci et al.: "Survey of machine learning techniques for malware analysis", Computers & Security. 81:123-147, 2019.
[2] Tamás et al.: "SIMBIoTA: Similarity-based malware detection on IoT devices", in Proc. of IoTBDS 2021, 58–69. SciTePress.
[3] Ferenc et al.: "Deep-water framework: The Swiss army knife of humans working with machine learning models", SoftwareX 12 (2020) 100551. Elsevier.

Please contact:

Levente Buttyán
Department of Networked Systems and Services, Budapest University of Technology and Economics, Hungary
buttyan@crysys.hu

Rudolf Ferenc
Department of Software Engineering, University of Szeged, Hungary
ferenc@inf.u-szeged.hu

Fighting Cybercrime by Introducing Trustworthiness and Interpretability in Deep Learning Malware Detection

by Giacomo Iadarola, Fabio Martinelli (IIT-CNR) and Francesco Mercaldo (University of Molise and IIT-CNR)

Cybercriminals can use a device compromised by malware for a plethora of purposes. Malicious intentions include the theft of confidential data, using the victim's computer to perform further criminal acts, or data ciphering to ask a ransom. Recently, deep learning is widely considered for malware detection. The main problem in the real-world adoption of these methods is due to their "black box" working mechanism i.e., the security analyst must trust the prediction without the possibility to understand the reason why an application is detected as malicious. In this article we discuss a malicious family detector, providing a mechanism aimed to assess the prediction trustworthiness and explainability. Real-world case studies are discussed to show the effectiveness of the proposed method.

In recent years, we have focused on cybercrimes perpetrated by malware that can compromise devices and IT infrastructures.. To fight cybercrimes due to malware spread, researchers are developing new methodologies for malware detection, with a great focus on the adoption of artificial intelligence techniques. One of the biggest controversies in the adoption of artificial intelligence models regards a lack of a framework for reasoning about failures and their potential effects. Governments and Industries are assigning critical tasks to artificial intelligence, but how can we ensure the safety and reliability of these models? In response to this need, the research community is moving toward interpretable models (the Explainable-AI field), able to provide the meaning of their predictions in understandable terms to humans [1]. One of the main reasons the adoption of these models in the real world is held back for effectively fighting cybercrimes exploiting malware is the lack of interpretation of the decision made by these models, which causes a lack of reliability certification, essential for achieving wide adoption.

We introduce a deep learning model for detecting malicious families in malware represented as images, with some interesting ideas for exploiting the activation maps, and provide model interpretability. We contextualise the proposed model on the Android platform, but the proposed method is platform independent.

The aim is to train a model i.e., a Convolutional Neural Network, for malware detection. For activation map drawing, we resort to the Grad-CAM [2] algorithm, able to generate a localisation map of the important regions in the image.

This mechanism provides to the security analyst a way to understand the reason why the model outputs a prediction. This aspect is fundamental for malware detection; in fact, malware belonging to the same family share parts of the code, hence areas of the images will be similar. For this reason, the activation map help to understand the area of the image symptomatic of a malicious behaviour exhibited by a family, thus providing trustworthiness.

We tested the model with a dataset composed of six malware families: Airpush, Dowgin, FakeInstaller, Fusob, Jisut and Mecor [3], obtaining a 0.98 accuracy.

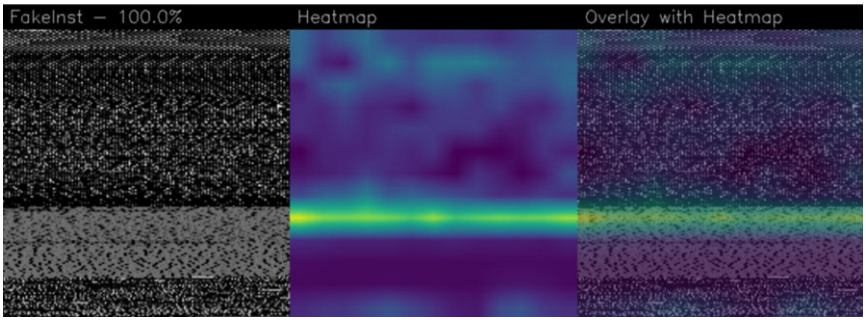


Figure 1: Sample identified by 7ab97a3d710e1e089e24c0c27496ee76 hash.

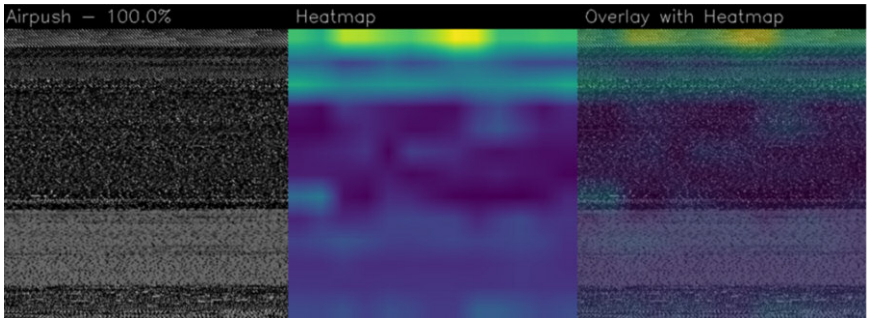


Figure 2: Sample identified by 7728411fbe86010a57fca9149fe8a69b hash.

To evaluate the classification trustworthiness, we exploited the activation maps. For each sample, we showed the greyscale image, the activation map and the picture generated from the overlap of the greyscale image and the activation map.

Figure 1 shows the activation map for a sample belonging to the FakeInstaller family, correctly detected.

The activation map exhibits an extended dark blue area (not of interest for the model) and several green and yellow areas (symptomatic of the activation).

By looking at the greyscale pictures in the left part of Figure 1, there is a characteristic grey area in the medium-low part of the images. The security analyst could make use of the activation maps to visualise which parts of the images were the most important for the classification.

Figure 2 shows the activation map of a sample belonging to the Airpush family. The greyscale image may look similar to the samples of the FakeInstaller family because of the grey area in the middle-low part of the image.

Figure 2 shows that the model focused attention on a different area of the malware with regard to the activation maps of the FakeInstaller sample; instead of focusing on the grey middle-low area, the activation map highlights the top area of the image, which seems to be the discriminating area for the Airpush family.

Artificial intelligence models are a powerful technology that will play a fundamental role in fighting cyber criminals. While the quantity of data to process keeps growing, artificial intelligence is becoming fundamental for protecting infrastructure and networks. Nevertheless, this technology suffers from bugs, inaccuracies and mistakes. We propose a malware detector by pointing out the importance of trustworthiness in deep learning. We suggest a technique whereby the security analyst is able to understand whether the prediction can be considered reliable. Future research plans will consider investigating the source code highlighted from the activation maps.

This work has been partially supported by EU E-CORRIDOR, EU SPARTA, EU CyberSANE project, and the RSE 2022 Cyber security and smart grids.

References:

[1] D. L. Parnas: "The real risks of artificial intelligence", Communications of the ACM, vol. 60, no. 10, pp. 27–31, 2017.

[2] R. R. Selvaraju, et al.: "Grad-cam: Visual explanations from deep networks via gradient-based localization", in Proc. of the IEEE international conference on computer vision, 2017, pp. 618–626.

[3] G. Iadarola, et al.: "Evaluating deep learning classification reliability in android malware family detection", 2020 IEEE Int. Symposium on Software Reliability Engineering Workshops (ISSREW), 2020, pp. 255-260.

Please contact:

Giacomo Iadarola, IIT-CNR, Pisa, Italy, giacomo.iadarola@iit.cnr.it

Fabio Martinelli, IIT-CNR, Pisa, Italy, fabio.martinelli@iit.cnr.it

Francesco Mercaldo, University of Molise & IIT-CNR, Campobasso, Italy, francesco.mercaldo@unimol.it, francesco.mercaldo@iit.cnr.it

Transparent and Explainable Information Quality Prediction

by Davide Ceolin (CWI)

Predicting the quality of the information online is a key step to contrast the spread of dis- and misinformation. Transparency and explainability of information quality prediction are key elements to increase their trustability and usefulness. We at CWI work on fostering online information quality explainability through transparent AI pipelines that combine argumentation reasoning, crowdsourcing, and logical reasoning.

The amount of information online and the impact it has on society imply the need for an automated prediction of information quality. However, different users in different contexts have very different needs and requirements. Therefore, these individuals can really benefit from the diversity of the information the Web provides. Some contexts might even require disinformation to be part of the picture, e.g., when researchers or journalists want to study, describe, analyse, or report on the nature of online disinformation itself. Therefore, the solution to the problem of disinformation, misinformation, and information excess is not to filter out low-quality information but to predict information quality to increase user awareness. As quality prediction can be perceived as subjective or biased, it is crucial that information prediction is both transparent and explainable. In other words, in order to increase the usefulness of information quality predictions, we need to win user trust in them (See Figure 1). Transparency and explainability are key ingredients to this aim. Also, explainability is a key ingredient to help identifying disinformation campaigns: by predicting diverse aspects of quality, we can unveil complex strategies that involve, for instance, the manipulation of narratives based on the combination of factual statements.

Transparency of information quality prediction guarantees that the computational steps performed to obtain quality predictions are accessible by humans so that they can follow the reasoning and understand how it has been implemented. This means that these computational steps should, ideally, implement well-known approaches from the humanities and social sciences, that are familiar to humans. For example, in our studies, we refer to argumentation theory and source criticism as useful theories on which to base our pipelines.

Computational argumentation reasoning is a vast field of AI that aims at studying how claims can be supported through diverse types of reasoning. We demonstrate that it is possible to implement argumentation reasoning through AI pipelines that combine natural language processing, machine learning, and logical reasoning, and use it to predict the quality of information items. For example, we test it on product reviews [1], and show that computational argumentation reasoning can be a useful tool to predict their quality in a similar manner as humans do. Here, we aim at transparency for the purpose of gaining actionable insights, so we are now working on understanding the implications of the choice of different implementations for the AI components involved in the pipelines (e.g., clustering algorithms, readability measures, etc.) in the information quality prediction performance.

Figure 1: Information quality prediction can be useful to users as long as labels are explainable and transparent to users (image source: flickr, "mysterious conspiracy" by ranma_tim, licensed under CC BY-ND 2.0.)

Source criticism is a well-known practice from the humanities, meant to determine the quality of information sources. Through specific checklists, scholars can determine whether a given book, journal, or article is of high enough quality to be considered as a source for their studies. We translated this practice in order to evaluate Web information items and again implemented this theory into transparent AI pipelines that use network analysis and evidential reasoning to help laypeople understand the quality of the information they consume online [2].

Information quality can be informally defined as 'fitness for purpose' and therefore, we can think of quality prediction as Boolean labels: information either fits a given purpose or not. However, such labels can be obtained in different manners. Transparency helps us understand the computational part of this, but then we need to understand also which aspects of quality are being considered. Information can, in fact, be more or less precise, neutral, complete, etc., and these aspects (or information quality dimensions) shed some light on

Figure 2: Transparent AI pipelines that combine natural language processing, machine learning, and logical reasoning are a tool for assessing online information quality (image source: www.pexels.com).

different and possibly independent characteristics of information. We can combine these predictions in order to understand whether a given information item fits a given purpose. However, when these aggregated predictions are presented to final users, it is important to explain them. Explaining information quality predictions means explaining which aspects of information quality were considered when predicting them. We performed experiments involving both experts and laypeople to this aim, using crowdsourcing platforms. These contributors were asked to evaluate statements and documents regarding the vaccination debate in one case, and regarding political statements in another one. Results show that we can obtain consistent results from human contributors [3].

These lines of research will be further investigated in the recently started Eye of the Beholder project [L1], which is led by Davide Ceolin and is a collaboration between CWI, the Netherlands eScience Center, and the University of Amsterdam. The project aims at extending an existing platform for transparent AI pipelines in order to provide media studies scholars with a tool to predict the quality of online information items in a transparent and explainable manner.

Link:

[L1] https://kwz.me/hjN

References:

[1] D. Ceolin, et al: "Assessing the Quality of Online Reviews Using Formal Argumentation Theory", Int. Conf. on Web Engineering (ICWE) 2021: 71-87.

[2] D. Ceolin, F. Doneda, G. Primiero: "Computable Trustworthiness Ranking of Medical Experts in Italy during the SARS-CoV-19 Pandemic", ACM 1st Int. Conf. on Information Technology for Social Good (GoodIT 2021): 271-276.

[3] M. Soprano, et al.: "The many dimensions of truthfulness: Crowdsourcing misinformation assessments on a multidimensional scale", Information Processing and Management 58(6): 102710 (2021), Elsevier.

Please contact:

Davide Ceolin, CWI, The Netherlands
Davide.Ceolin@cwi.nl

SPOTTED: Systematic Mapping of Detection Approaches on Data Sources for Enhanced Cyber Defence

by Manuel Kern and Florian Skopik

In the last decade there was a clear paradigm shift from focusing only on prevention and protection to also including detection and response. While prevention and protection are indispensable to enable a baseline security, it is presumed that attackers have already compromised systems to some extent ("presumption of compromise"). The fact that professional attackers often operate in the network over a long period of time has long been known in cyber security research. A key pillar of a holistic security approach is therefore the early detection of attackers in the network. But still, the average time to detect attackers remains high. In the course of a study commissioned by IBM Security [L1], the average time it takes to detect a data breach is quantified with a time period of 212 days, five days longer than the year before.

It was in late 2020 that the disastrous case of SolarWinds became public [L2]. State attackers abused the update mechanism of a security solution to infiltrate thousands of organisations up to the highest state level and to move unnoticed in the networks of the attacked organisations for many months. Kaspersky [L3] reports that compared to costs of a cyberattack with immediate remediation, recovery costs are four times higher if remediation is per-

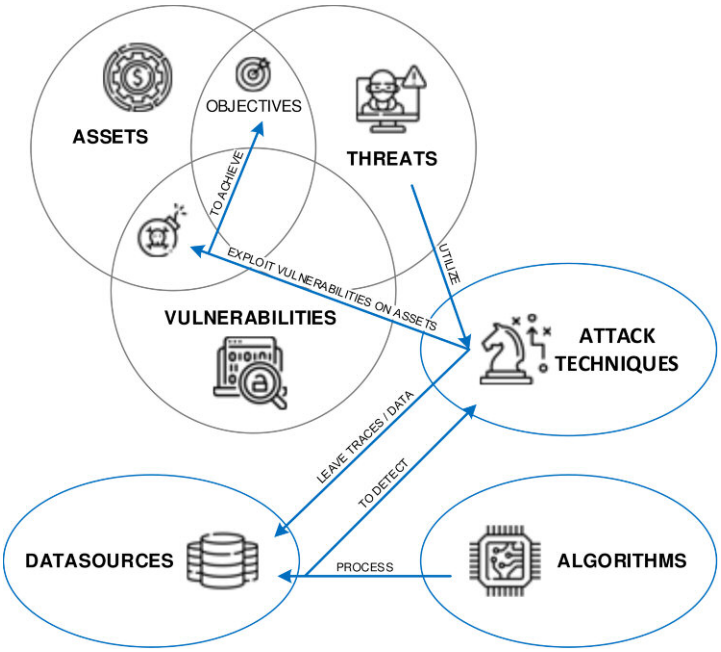


Figure 1: D3TECT's attack detection in a nutshell.

formed after one week. At the same time, detection within one day is still about 30% more cost-effective than detection after more than one week. The insurance provider Allianz [L4] has determined a total loss of 660 million euros in an evaluation of 1,736 incidents. The largest share of costs is due to operational downtime. Another emerging threat is theft of customer data and blackmailing.

Implementing organisation-wide detection and response to deal with this issue, is a resource-intensive undertaking. Not only are required software solutions are difficult to select, deploy, and maintain, it also requires expensive and often rare security expertise of staff to install, maintain and operate these solutions. Security experts are a highly requested resource these days and at the same time, trends driven by economy, environment and technology have drastically accelerated digitalisation. This is also reflected in current employment figures, which reflect a clear lack of IT and IT security specialists worldwide. Besides the lack of human resources, detection and response needs dedicated infrastructures with excessive performance requirements. Efficient detection systems for large infrastructures are not an off-the shelf product. The security aspects of ongoing system integration are complex and typically not fully considered in business decisions. Implementing infrastructure-wide detection systems can quickly consume the entire IT security budget. There is a high probability that detection and response projects will be rejected from the beginning, aborted, or carried out incompletely. This is indeed a serious problem. Thus, it requires effective detection and response to reduce the time of cyberattacks and to keep economic damage and impacts on human safety at a minimum.

The mission of project SPOTTED is to counteract these problems by lowering the entry barrier for modern monitoring and detection solutions and making them more applicable. In recent years, a wide variety of novel methods and models for incident detection and response have been developed that enable accurate detection and efficient prediction in minimum time. Special focus is put on the fact that organisations have only limited resources to establish and operate them. Thus, an optimisation problem is the basis of the project. Cyberattacks leave traces in data sources, such as in log files, memory or data-streams. Detection systems utilise these data sources to detect the application of specific attack techniques. Attack techniques vary considerably in terms of their effectiveness, potential impact and application by threat actors. Data sources, on the other side, may contain traces of one or several attack techniques, and the effort to process their output may differ heavily. Therefore, it is obvious that not all data sources are of equal value for detection and organisations must carefully survey which sources shall be analysed and what attack techniques need to be discovered.

SPOTTED developed D3TECT, a process model based on the three key elements of attack detection (techniques, data sources and algorithms). Figure 1 outlines interactions of key elements on an organisation's assets, their vulnerabilities and threats. The model describes a procedure for dynamically ranking and selecting data sources suitable for detection. The novelty is that this model accounts for constraints in the selection process. For instance, if a certain data source cannot be utilised in a specific setting, e.g., due to data privacy constraints, the discovery of the most important attack techniques is still ensured by the remaining data sources. Eventually, the D3TECT approach solves the challenge of strategically selecting data sources while accounting for their varying usefulness for attack detection. The model is tested with real data, utilising the MITRE ATT&CK framework and numerous public cyber threat intelligence databases. A recent work [1] shows the ranking results and discusses their plausibility to validate D3TECT.

About the project

The project SPOTTED is financially supported by the Austrian Research Promotion Agency (FFG) under grant number FO999887725). The project is carried out in the course of an industry-related PhD thesis at the Austrian Institute of Technology (AIT) in cooperation with the Vienna University of Technology (TU WIEN).

Links:

[L1] https://kwz.me/hff
[L2] https://kwz.me/hfj
[L3] https://kwz.me/hfq
[L4] https://kwz.me/hfv

Reference:

[1] M. Kern, et al.: "Strategic selection of data sources for cyber attack detection in enterprise networks: A survey and approach", 37th ACM/SIGAPP Symposium On Applied Computing, 2022.

Please contact:

Manuel Kern, AIT Austrian Institute of Technology, Austria
manuel.kern@ait.ac.at

Kyoushi Testbed Environment: A Model-driven Simulation Framework to Generate Open Log Data Sets for Security Evaluations

by Max Landauer, Florian Skopik, Markus Wurzenberger and Wolfgang Hotwagner (AIT)

Cyber security leverages intrusion detection systems that analyse log data and network traffic to disclose suspicious activities and protect networks against cyberattacks. Verifying the functionality and measuring the effectiveness of these detection systems is not trivial, since it usually is not desirable to launch actual attacks in an organisation's productive infrastructure. Therefore, such evaluations are often carried out in isolated testbeds, i.e., simulated networks comprising components and applications that are representative of their real-world counterparts in terms of configuration, scale, and utilisation. However, setting up and maintaining such testbeds is complex and labour-intensive, particularly when experiments are required to be reproducible and adaptable. To alleviate these issues, we developed the Kyoushi Testbed Environment, an open-source simulation framework that enables automatic and parallel testbed instantiation through model-driven design, simulation of normal user activities to generate a baseline workload, injection of attack scenarios with variations, and labelling of collected log data.

Despite a great need, there are hardly any publicly available log data sets that are suitable for security exercises, such as evaluations of attack detection and classification solutions. The main problems with existing data sets are: outdated or oversimplified use-cases, processed or anonymised logs, incomplete documentations, and missing reproducibility. Understandably, organisations are reluctant to make log data collected at their premises publicly available, as they likely contain traces of sensitive information, such as usernames, network configurations, setup structures, asset information, or software versions [1]. In addition, adversaries could possibly gain insights on deployed solutions and configurations from log data and target them in attacks.

Testbeds do not suffer from such issues, as they are isolated from the production environment and therefore allow the launching of attacks against services without the fear of any adverse consequences. Another advantage is the fact that simulated normal activities are clearly discernible from attack manifestations as all activities that are expected to occur are known beforehand. This facilitates generation of a ground truth table that specifies attack times and malicious events, and is essential for computing detection accuracies in evaluations. On top of that, analysts have full control over all settings of the simulation running on a testbed, meaning that they can arbitrarily adjust simulation parameters such as the network size or average utilisation of services [2].

To further ease and automatise the process of adapting the simulations, the Kyoushi Testbed Environment incorpo-

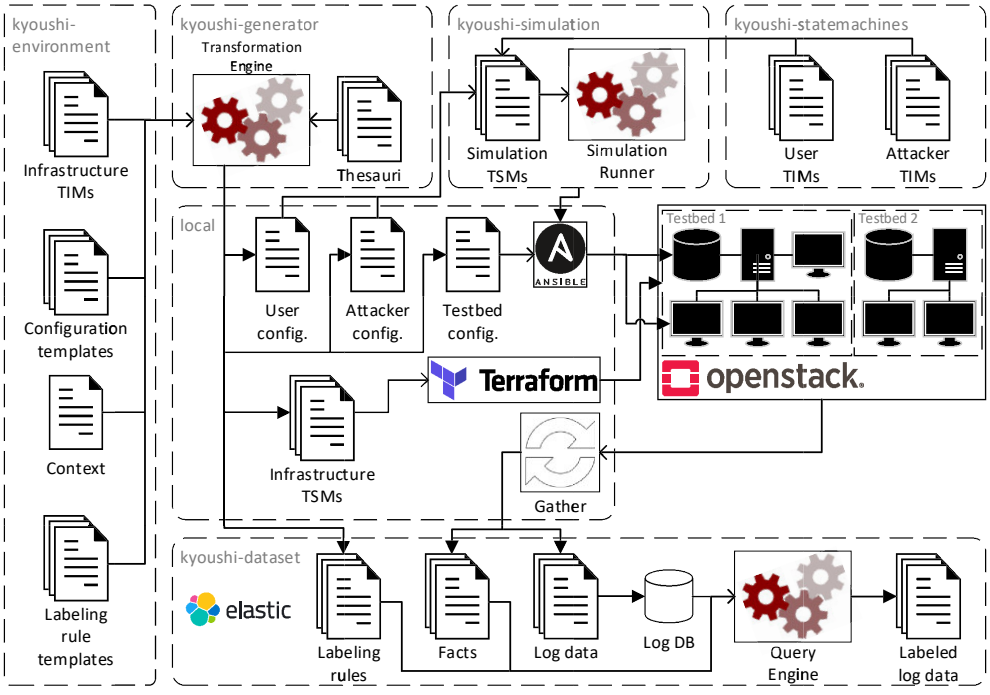


Figure 1: Overview of the Kyoushi Testbed Environment components and log data generation process.

rates concepts from model-driven engineering that select relevant testbed parameters from predefined dictionaries or distributions, for example, usernames are randomly chosen from thesauri and executed activities are randomly selected based on probability distributions. Designing testbeds from such abstract models introduces variations in the resulting log traces, which is advantageous for several reasons: (i) it facilitates log collection of an arbitrary number of testbeds that represent different technical environments, (ii) it increases robustness of results when evaluating intrusion detection systems, and (iii) it uses repeated executions of similar attacks in evaluations of alert aggregation approaches [3].

Technical Overview of the Kyoushi Testbed Environment

The Kyoushi Testbed Environment is a modular framework for testbed generation, behaviour simulation, and data handling. Figure 1 shows a conceptual overview of all involved components. The left side depicts the kyoushi-environment, which is the main component that defines the scope of the testbed and provides all models required to set up the technical infrastructure of the simulated network. Models are templated scripts that do not specify testbed parameters, such as usernames and IP addresses, and are therefore referred to as testbed-independent models (TIM). The kyoushi-generator then ingests these models and fills out all missing information. The resulting scripts are stored in the local environment and allow deployment of a specific testbed instance, and are accordingly referred to as testbed-specific models (TSM). Provisioning tools, such as Terraform, are then capable of creating the network and machines on virtualisation platforms, such as OpenStack, in a fully automatic process.

In addition to hardware provisioning scripts, the configurations of the user simulation and testbed are generated as part of the transformation of TIMs to TSMs. The testbed configuration comprises setup and initialisation scripts for services and applications, for example, databases and content management systems, that are suitable for automatic deployment with software provisioning tools, such as Ansible.

To generate a baseline of normal workload on the network, the kyoushi-statemachines module provides TIMs for all possible activities carried out by users and attackers, such as writing emails, browsing the Internet, executing commands, etc. In the kyoushi-simulation component, these state machines are combined with configurations that specify parameters, such as state transition probabilities, to yield simulation TSMs. These TSMs are executed by the simulation runner, e.g., a web automation framework such as Selenium.

The simulation may be stopped at any desired time, typically after several hours or days. A script is used to gather various log files from all hosts in the testbed, including authentication logs, access logs, error logs, application logs, syslog, network traffic, etc. The kyoushi-dataset module then stores the logs in a database and uses labelling rules generated alongside the TSMs to identify log events related to attacks.

The Kyoushi Testbed environment was used to instantiate eight enterprise IT networks comprising web servers, cloud shares, groupware, etc., that vary in terms of network size, configuration, and utilisation. As part of the simulation, several attacks, such as security scans, data exfiltration, exploits, and password cracking, were launched against the servers. The data collected from these testbeds were labelled as suitable for forensic security evaluations, and made available open source [L1, L2].

Links:

[L1] https://zenodo.org/record/5789064
[L2] https://kwz.me/hf0

References:

[1] R. Uetz, et al.: "Reproducible and Adaptable Log Data Generation for Sound Cybersecurity Experiments", in Proc. of the Annual Computer Security Applications Conf., pp. 690-705. ACM, 2021.

[2] F. Skopik, et al.: "Semi-synthetic Data Set Generation for Security Software Evaluation", in Proc. of the Annual Int. Conf. on Privacy, Security and Trust, pp. 156-163. IEEE, 2014.

[3] M. Landauer, et al.: "Have It Your Way: Generating Customized Log Data Sets with a Model-driven Simulation Testbed", IEEE Transactions on Reliability, Vol.70, Issue 1, pp. 402-415. IEEE, 2021.

Please contact:

Max Landauer, AIT Austrian Institute of Technology, Austria
max.landauer@ait.ac.at
+43 664 88256012

A Scalable Ensemble-based Framework to Analyse Users' Digital Footprints for Cybersecurity

by Gianluigi Folino, Francesco Sergio Pisani (ICAR-CNR). and Carla Otranto Godano (HFactor Security)

In the field of cybersecurity, it is of great interest to analyse user logs in order to prevent data breach issues caused by user behaviour (human factor). A scalable framework based on the Elastic Stack (ELK) to process and store log data coming from digital footprints of different users and from applications is proposed. The system exploits the scalable architecture of ELK by running on top of a Kubernetes platform, and adopts ensemble-based machine learning algorithms to classify user behaviour and to eventually detect anomalies in behaviour.

In recent years, the number of cybersecurity attacks has been increasing and generating ever-larger amounts of data, often compromising computer networks by exploiting user weaknesses. An average organisation is targeted by over 700 social engineering attacks in one year, due to human behaviour [L1]. Often, the detection of these behaviours

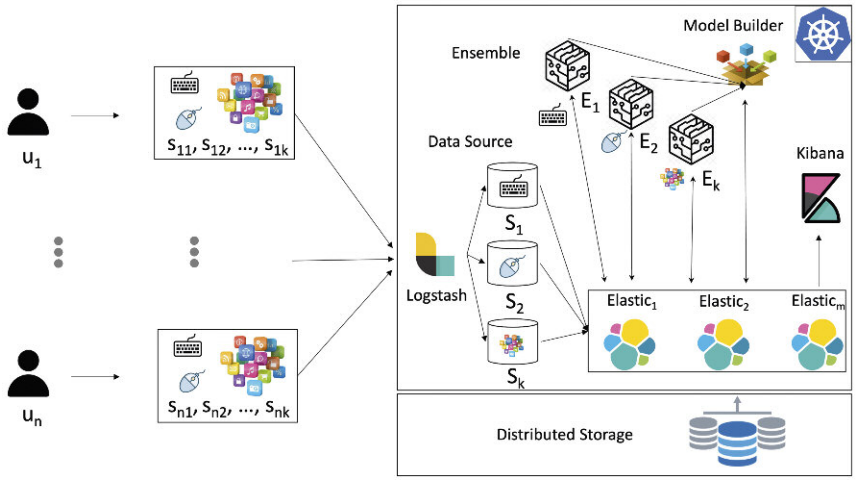


Figure 1: The software architecture of the Kubernetes-based ELK cluster.

or of the consequent vulnerabilities happen when the attack has already occurred; on the contrary, a proactive solution, which prevents vulnerabilities from being enabled, is necessary. In addition, in order to operate with the security weaknesses derived from the human factor, a number of critical aspects, such as profiling users for better and more focused actions, analysing large logs in real-time, and working efficiently in the case of missing data, are needed.

According to an industrial report [L2], phishing continues to be the top threat action used in successful breaches (it nearly doubled in frequency from 2019 to 2020) linked to social engineering, with login credentials stolen in 85% of these breaches. However, a security awareness training program can significantly reduce these risks (i.e., from 31.4% to 16.4%, with a limited training period of three months). Analysing user behaviour, to identify homogenous categories or for detecting anomalies, can be a winning strategy to plan a focused training program. However, this task requires processing very large datasets and the sources of user behaviour are heterogeneous and can be missing.

To cope with these issues, a joint collaboration between ICAR-CNR and a cybersecurity startup, HFactor Security [L3], which supplied the ELK-based platform and the data coming from real users, permitted the design of a scalable framework [1]. The framework adopts a distributed ensemble-based evolutionary algorithm [2] to classify the user behaviour and to eventually detect anomalies in their behaviour. The framework (i) operates even with missing sources of data, (ii) works in an incremental way and with streaming data, and (iii) uses a scalable ELK architecture for processing real, large logs.

In more detail, to analyse the user behaviour, three main sources of data are processed: keyboard usage (storing only the zones of the keyboards for privacy reasons), mouse usage (considering only the subarea of the screen in which the user clicks or moves the mouse), and the main application/categories in which the user spends time. These data can be stored in an efficient and scalable architecture, based on a kubernetes ELK cluster, illustrated in Figure 1.

On the left of this architecture, a number of agents (installed on the PCs of the users) collect information concerning different sources of information and/or monitoring tools (i.e., generally automatic software analysing the actions and the behaviours of the users). These data are collected and usually stored in log files, which are sent to the cluster by using Filebeat agents.

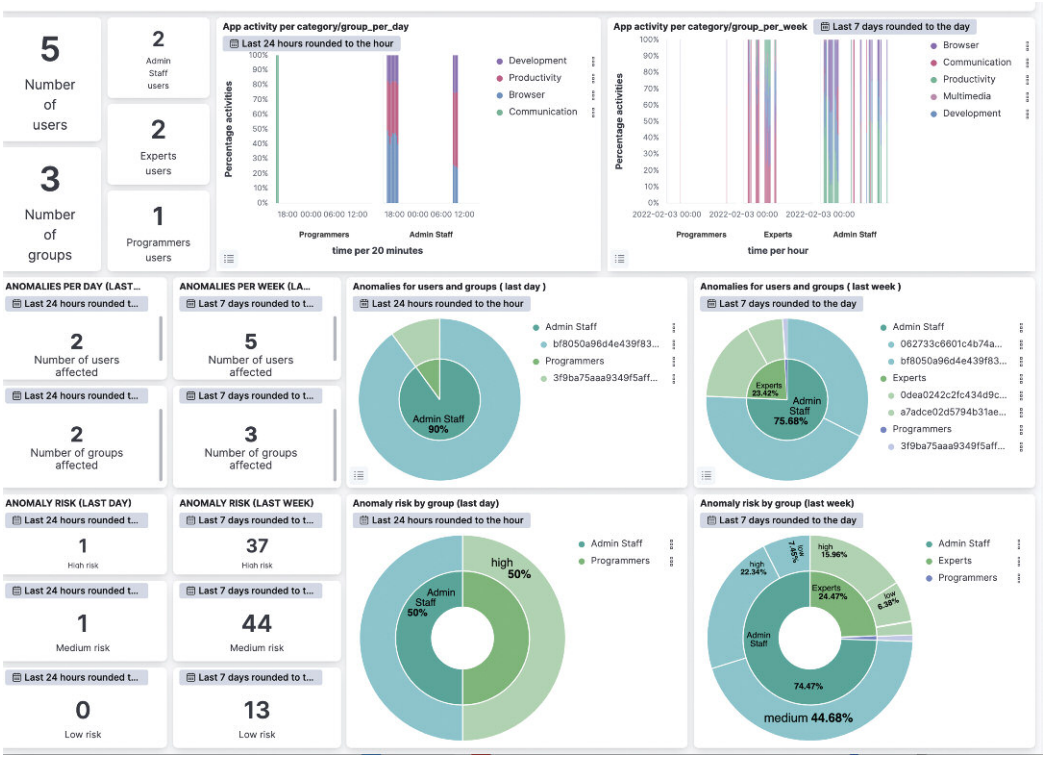


Figure 2: Kibana overview dashboard (anomalies, and class of risk for groups and users).

On the right side of the figure is illustrated the ELK cluster architecture running on top of a Kubernetes platform. The components are Kubernetes Service and they can be replicated to fulfil load demands. The Logstash Service processes the data coming from the Beats agents by applying some specific filtering and transformation rules. Note that the multi-source logs are gathered by Logstash and opportunely grouped into homogenous types (i.e., the data concerning the output of the keyboard or the process usage of the different users). Then, the data are transmitted to the Elastic services that store them on the elastic indexes placed on a persistent volume, managed with a distributed file system.

When these data are stored, the Kibana service and the Model Builder components can have access to this information for visualisation and computation. The flexibility and user-friendly Kibana online data visualisation platform is designed to serve as an end-point for querying and access the amount of user-mined data gathered through the Elasticsearch indexes. Indeed, the latter is responsible of the machine learning tasks, and of ingesting and storing the large-scale data in an automated manner. In addition, the information extracted from the real user logs will be stored for further analysis, i.e., the classification task and the detection in real-time of suspect behaviours and anomalies.

The ML component of the architecture is capable of training and executing classification models to detect anomaly behaviours. The goal is to detect extraneous activities of users belonging to pre-defined group, each one defined as a cluster of users that exhibit the same behaviours. The ML models enrich the collected data, helping the operators to detect and mitigate risk in their organisation. The anomaly detection task was reformulated as a combination of user/group identification tasks, following the principle that the lower the probability that a digital footprint belongs to its corresponding user/group, the more its behaviour is anomalous.

Finally, by using Kibana, the information about these anomalies is processed and plotted with several histograms and charts. In Figure 2, the dashboard showing an example of visualisation of the anomalies, user and group, and the class of risk for a specific time range defined by the Data Protection Officer (DPO) is shown.

Experimental results, conducted using a benchmark dataset and real users, show that the framework is effective, even in the case of missing data, in classifying the behaviour of different users and in detecting the anomalies inside the user/group behaviour, with a low number of false alarms. For future works, the solution will be tested on a real scenario with logs coming from many hundreds of real users.

Links:

[L1] https://kwz.me/hfd
[L2] https://kwz.me/hfc
[L3] http://www.hfactorsec.com/

References:

[1] G. Folino, C. Otranto Godano, F. S. Pisani: "A Scalable Architecture Exploiting Elastic Stack and Meta Ensemble of Classifiers for Profiling User Behaviour", 30th Euromicro Int. Conf. on Parallel, Distributed, and Network-Based Processing, IEEE, 2022.

[2] G. Folino, F. S. Pisani: "Evolving Meta-Ensemble of Classifiers for Handling Incomplete and Unbalanced Datasets in the Cyber security Domain", Applied Soft Computing, Elsevier, pp. 179-190, Vol. 47, October 2016.

Please contact:

Francesco Sergio Pisani and Gianluigi Folino, ICAR-CNR, Italy
francescosergio.pisani@icar.cnr.it, gianluigi.folino@icar.cnr.it

Timestamp Patterns in Windows Forensics

by Robert Luh (University of Vienna and St. Pölten University of Applied Sciences) and Michael Galhuber (St. Pölten University of Applied Sciences)

Timestamps are among the most expressive artefacts in a digital forensic investigation. Our research shows that the distinct patterns caused by the interaction with individual files can yield more insight than previously documented and enables application fingerprinting within a Windows environment through timestamps alone. Furthermore, we classify timestamp forgery tools and present a means to detect their use.

Forensics practitioners heavily rely on timestamps in their investigation of criminal cases to create a meaningful timeline of events. Modern filesystems, such as Microsoft's NTFS, track the creation, modification, access, and metadata changes down to near nanosecond resolution [1]. In addition to an exposed and easily accessible set of timestamps, NTFS maintains a second set that is typically altered only by the Windows kernel. These two sets (see figure), which are stored as attributes in the Master File Table (MFT) for each file created on the system, do not necessarily hold the same information but generally complement each other in terms of relevance. The inherent rules specifying when certain changes to these 8 timestamps are committed to the disk, result in unique patterns that potentially offer in-depth insight into a user's or program's activity that goes beyond other forensic data sources.

Experimental research [2] at St. Pölten University of Applied Sciences and the University of Vienna has shown that these patterns are even more expressive than previously documented [L1]. While it has been established that it is possible to identify generic file creation, access, modification, renaming,

copying, moving, and deleting operations through NTFS timestamps, we have gone a step further and investigated the means by which these operations were conducted. We identified patterns that allow investigators to determine which specific application was used when interacting with a file, or how a user triggered a certain operation. For example, we were able to identify which specific PDF printer was used to create a document; if a file was opened in a certain text editor or on the command line; whether a file was copied or moved using a PowerShell command, keyboard shortcut or context menu (right click); or which format was used for a newly saved picture file, without assuming the validity of the extension.

In the second part of the project, we used that self-same approach to spot timestamp forgery, which constitutes one of the most disruptive anti-forensic techniques [3]. In all cases but one, we could determine that timestamps had been altered and were often able to identify the specific tool.

The practical implications in regard to these findings are manifold. Forensic investigators can use our timestamp patterns and updated time rules to correlate their findings by, e.g., linking an individual file to an application present on a suspect machine, even if additional app usage artefacts are missing or have been deleted. User profiling through shell interaction becomes possible, as does identifying the use of (shell) scripts for automated file modification performed by malicious software. Some file types may even be recognised through their timestamps after their extension and file signature (magic number) have been removed or falsified.

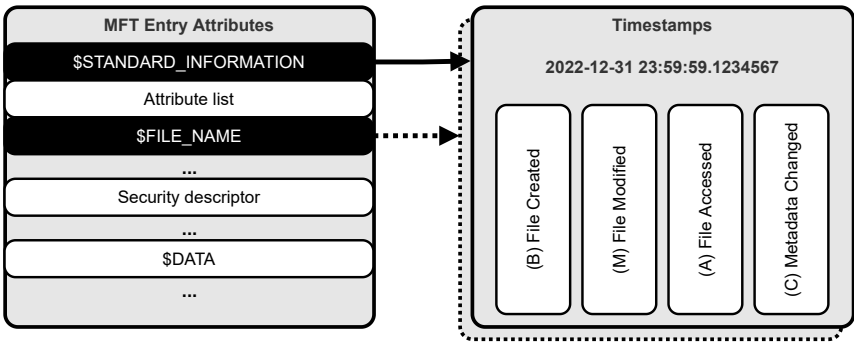


Figure 1: Excerpt from an MFT file record with the two highlighted attributes containing four timestamps each.

With the identification and classification of forgery tools by their impact and level of access (ranging from common API functions to direct disk access), we hope to make investigations more resilient to anti-forensics techniques, and help analysts quickly determine which artefacts can be trusted and which should be examined further.

In future research, we will endeavour to train machine learning models to automatically identify a wider range of applications, and develop a framework for forensic scientists to build their own models tailored to specific use cases and application sets. The goal going forward will be to make the investigative process more efficient and focus on combating anti-forensic techniques, which are often overlooked due to tool and time constraints.

Link:
[L1] https://kwz.me/hfy

References:
[1] B. Carrier: "File System Forensic Analysis", Addison-Wesley Professional, 2010.
[2] M. Galhuber and R. Luh: "Time for Truth: Forensic Analysis of NTFS Timestamps" in ARES WSDF, 2021.
[3] W. Minnaard, et al.: "Timestomping NTFS", MSc research project, University of Amsterdam, 2014.

Please contact:
Robert Luh
University of Vienna and St. Pölten University of Applied Sciences, Austria
robert.luh@fhstp.ac.at

Michael Galhuber
St. Pölten University of Applied Sciences, Austria
is211816@fhstp.ac.at

Meta-framework for Automating Static Malware Analysis

by Patrick Kochberger, Sebastian Schrittwieser (University of Vienna) and Edgar R. Weippl (SBA Research)

In cybercrime, malware plays a weighty role and malware authors heavily rely on different code obfuscation techniques such as packing, virtualisation, or control flow transformations, and other anti-analysis methods to hide malicious functionality in binary code. With thousands of new malware samples emerging every day, efficient analysis is crucial for fighting malware-based cybercrime. We present a novel meta-framework for malware analysis that helps find the optimal analysis strategy for a malware sample. The research for the work was conducted in a joint project together with the University of Gent in Belgium [L1].

Code obfuscation [1] is widely used for protecting benign software, but also for hiding malicious functionality in malware. The basic idea of obfuscation is to intentionally modify code in such a way that its (malicious) functionality is more difficult to detect, and analysis becomes more time-consuming. In malware identification, potentially malicious code samples are often analysed in dynamic malware sandboxes, which observe their functionality and interaction

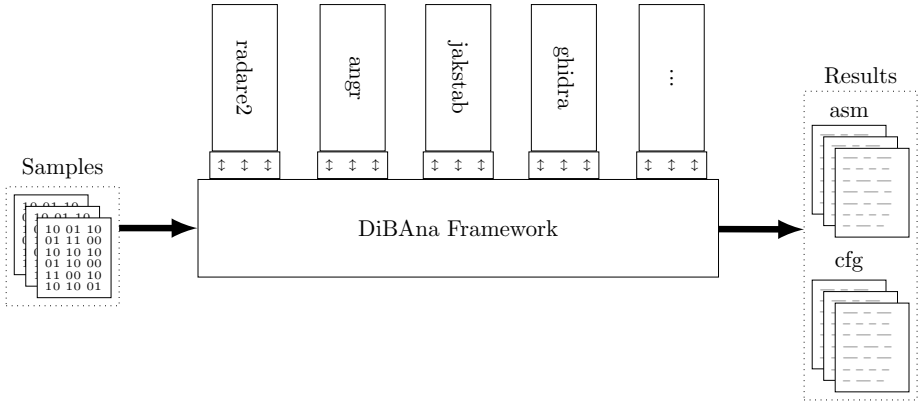


Figure 1: A fully automatic, novel meta-framework.

with the operating system at runtime. Malware, however, often can find out that it is running in a sandboxed environment and then stops its malicious activities to avoid detection.

The second important malware analysis methodology is automated static code analysis to get a quick insight into the functionality that is contained in a binary. The state-of-the-art in static code analysis has made great strides in speed and coverage in recent years [2]. Generally speaking, static code analysis aims at approximating a program's behaviour without actually executing it. A static analysis returns only incomplete information from which assumptions have to be made. Thus, only incomplete approximations can be made from analysing a program statically and the results of different analysis tools will differ. Especially for obfuscated binaries, this means that multiple static analysis tools can result in very different approximations, and while some tools might deliver usable results, others might fail completely. In addition, the static code analysis landscape is highly diverse with methodologies ranging from formal model checking over data flow analysis to machine-learning-based approaches. The results from all these different methodologies will highly depend on the analysed binaries, their structure and applied code obfuscations.

To be able to efficiently find the optimal static code analysis methodology for a given malware binary, we developed a fully automatic, novel meta-framework that runs multiple analyses in parallel and compares its results.

The framework consists of so-called modules and actions (see Figure 1). Actions represent reverse engineering or analysis tasks (e.g., listing functions, disassembling bytes, reconstructing a control flow graph, etc.). The tool-specific modules translate the actions selected in the configuration of an analysis run into the specific parameters required for the various binary analysis frameworks. In our implementation of the framework, a module consists of a simple shell script for setup and a Python class for the actual translation tasks, which derives from a generic analysis base class. For each framework, the class connects to its API, calls the individual tools and collects the output for a certain task. The output is then cleaned, normalised, and can be used for further analysis tasks or comparison with the results from other tools.

The binary analysis frameworks that we integrated in the meta-framework are freely available, not cloud-based, and provide an API or scriptable interface of some kind. Currently, the framework includes the static binary analysis tools radare2, rizin, angr, AMOCO, BARF, Capstone, Distorm3, Ghidra, objdump and Jakstab. Besides basic information on a sample (MIME-type, extension, architecture, etc.), the framework allows extracting the control flow graph, functions, sections, and the disassembly of the code.

The introduced meta-framework for static binary analysis is a first step into making large-scale malware analysis more efficient, as its results indicate the most promising methodology for further analysis tasks to a human analyst. In the future, we aim to use the framework for collecting large-scale datasets of analysis runs from different types of binaries (e.g., built using different compilers and obfuscations). With the help of machine-learning, we then want to identify which types of binaries are best analysed with which static code analysis methodology – even before running one of the analysis tools.

The project EMRESS [L2] is funded by the Austrian Science Fund (FWF) under grant I 3646-N31. The financial support by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

Links:

[L1] https://www.ugent.be/ea/elis/en/csl

[L2] https://kwz.me/hfw

References:

[1] S. Schrittwieser et al.: "Protecting Software through Obfuscation: Can It Keep Pace with Progress in Code Analysis?", ACM Computing Surveys 2017. https://doi.org/10.1145/2886012

[2] C. Pang et al.: "SoK: All You Ever Wanted to Know About x86/x64 Binary Disassembly But Were Afraid to Ask", IEEE SP 2021. https://doi.org/10.1109/SP40001.2021.00012

Please contact:

Patrick Kochberger
University of Vienna, Austria
patrick.kochberger@univie.ac.at

Sebastian Schrittwieser
University of Vienna, Austria
sebastian.schrittwieser@univie.ac.at

Edgar R. Weippl
SBA Research, Austria
EWeippl@sba-research.org

ARIMA Security Metrics: Facilitating Decision-making Processes and Situational Awareness in Threat Intelligence

by Jan Kohlrausch (DFN-CERT)

At DFN-CERT, we work on augmenting Security Metrics with a family of stochastic models. For a given Security Metric, an Autoregressive Integrated Moving Average (ARIMA) model is selected that encapsulates the sequence of metrics results and provides objective mathematical properties. This additional mathematical layer results in a better understanding of the metrics properties, facilitates decision-making processes, and supports situational awareness in Threat Intelligence.

Security Metrics and Threat Intelligence

Following the National Institute of Standards and Technology (NIST), Threat Intelligence and Security Metrics can be briefly summarised as follows:

- Threat Intelligence can be characterised as refined threat information supporting decision-making processes ([L1]). Moreover, NIST defines threat information as any information that can help an organisation to identify, assess, monitor, and respond to Cyber threats (e.g., Indicators of Compromise).
- Security Metrics are a tool to facilitate decision-making processes ([L2]). Technically, metrics usually quantify or measure security data. For example, an important efficiency metric in incident response is the "mean-time-to-fix" recovering from an attack.

Building upon the common objective of facilitating decision-making processes, Security Metrics are a promising tool implementing the refinement process for Threat Intelligence.

Augmenting Security Metrics with ARIMA Models

Considerable efforts have been spent on Security Metrics, addressing the design of metrics (e.g., "SMART" criteria) and their classification into different groups. In practice, a large number of metrics for Information Security has been proposed. In almost all of these metrics, the underlying technical result is a sequence of time-dependent measurements (aka time series). According to the definition, this sequence has to facilitate or drive the decision-making process of the corresponding metric. Our key finding is that this process benefits by mathematically modelling the metrics sequence of results, providing additional objective data properties. We selected ARIMA models because they are a solid and well-researched approach stochastically and have already proved their applicability for many similar use cases. From an operational point of view, the ARIMA data model provides an additional mathematical layer that improves a Security Metric by the following properties:

- Estimating measurement and data uncertainties: ARIMA models allows estimation of the uncertainties and variances affecting the measurements of a Security Metric. It is important to note, that this especially works for random errors; however, systematic errors have to be addressed by other means.
- Predicting metrics values and detecting anomalies: Upcoming results and their margin of uncertainty are predicted by ARIMA models, allowing the deduction that a new value does not fit into the series of historic results. Such anomalies (mathematically outliers) are all measured values that are outside of the predicted uncertainty margin (confidence interval). Since anomalies may be caused by malicious activity pertaining to a severe incident, they require a further analysis to identify their root cause. Technically, this could be addressed by putting additional sophisticated sensors in place, providing data for a more detailed analysis of the attack activity.
- Facilitating decision-making processes based on objective criteria: The ARIMA approach provides a detailed model of the metrics results. In addition to anomaly detection, time series analysis provides information whether the results are statistically constant (stationary stochastic process) or if there is a constant change that can be caused by a linear trend. Detecting trends is important, because they indicate a changing threat landscape, which may have a considerable impact on security governance. Thus, these mathematical properties are of great importance and complement or may even substitute an interpretation based on gut feeling or visual inspection.

The process of augmenting Security Metrics with ARIMA models is further detailed in [1].

For the previously mentioned "mean-time-to-fix" metric, an ARIMA model substitutes the simple “mean” by a mathematical model that accounts for more complex data properties. As detailed above, the model allows the prediction of the time-to-fix for future incidents, which has a serious impact if incidents grow in complexity or severity over time.

Application for Situational Awareness in Threat Intelligence

Because of the capabilities of ARIMA Security Metrics detecting anomalies, situational awareness in Threat Intelligence becomes a promising field of application. In the following, we demonstrate how an ARIMA Security Metric is applied to detect a significant increase in attack activity against the Microsoft Azure cloud service pertaining to the "OMIGOD" exploit. It is based on the report and data of the Internet Storm Center (ISC) of the SANS institute that was published on 20 September 2021 [L3].

The ISC data that represent the results of the ARIMA Metric “Daily number of targets being attacked on port TCP/12702 are shown in Figure 1 (blue line). A rapid rise of attacks can be seen starting on 15 September 2021. For fitting the ARIMA model, we used the Python module "statsmodels". Based on the Box-Jenkins method, the ARIMA (2,0,2) model has been selected and ap-

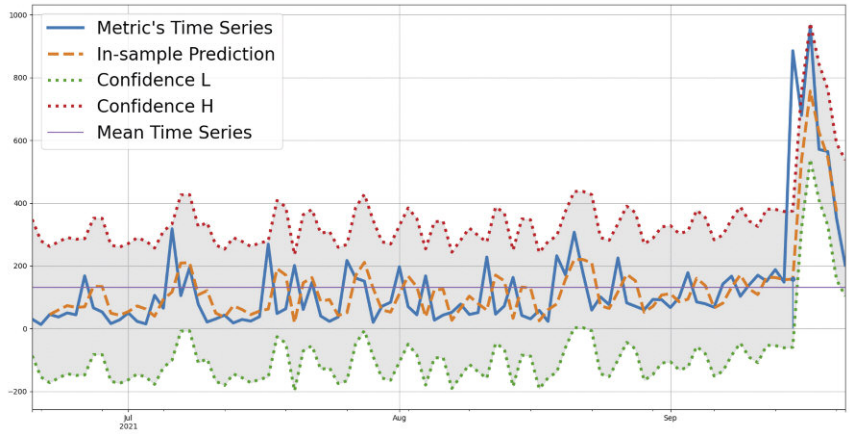


Figure 1: Number of targets being attacked on port TCP/1270 (ISC data pertaining to OMIGOD report [L3]) and ARIMA analysis (in-sample prediction and confidence interval).

plied to predict in-sample data points (orange dashed line) and the 95% confidence intervals of the prediction (gray area between red and green dotted lines). It can be seen that the ARIMA Security Metric reliably identified the rapid rise in attack activity on 15 September 2021 (the measured value exceeds the 95% confidence interval by a lot). Although the rise can be easily spotted, manual inspections or interpretations do not scale for a larger number of graphs. In contrast, ARIMA Security Metrics are an ideal tool for automatically monitoring a virtually unlimited number of time series to detect unusual activity requiring a deeper analysis. Thus, ARIMA Security Metrics can direct the attention to the critical events or situations that require further analysis.

ARIMA Security Metrics are currently deployed in the CONCORDIA Horizon 2020 project [L4] delivering attack landscape awareness for the data in the Treat Intelligence platform.

Links:
[L1] https://kwz.me/hfx
[L2] https://kwz.me/hfz
[L3] https://kwz.me/hfA
[L4] https://www.concordia-h2020.eu/

Reference:
[1] J. Kohlrausch, E. A. Brin: "ARIMA Supplemented Security Metrics for Quality Assurance and Situational Awareness", in Digital Threats: Research and Practice, Volume 1, Issue 1, March 2020, https://doi.org/10.1145/3376926

Please contact:
Jan Kohlrausch, DFN-CERT, Germany
kohlrausch@dfn-cert.de

From Collaboration to Automation: A Proof of Concept for Improved Incident Response

by Lasse Nitz (Fraunhofer FIT), Martin Zadnik (CESNET), Mehdi Akbari Gurabi (Fraunhofer FIT), Mischa Obrecht (Dreamlab Technologies AG) and Avikarsha Mandal (Fraunhofer FIT)

Effective incident response relies on taking accurate and timely measures in reaction to cybersecurity incidents. The increase in both the number and variety of cyberattacks, however, makes it challenging for incident handlers to keep up with this task. In the H2020 project SAPPAN, we take a practical look at this problem and explore the sharing of incident handling information, the automation of incident response processes, as well as the relationship between these two topics, to assist human operators in their work.

Automation within and information sharing between computer security incident response teams (CSIRTs) have the potential to improve response times for both common and novel attacks, despite a seemingly ever-increasing number of cybersecurity incidents. High-quality detection systems and automation of common incident response processes can help CSIRTs to utilise the time and efforts of operators more effectively by allowing human experts to focus on critical and novel kinds of attacks. Sharing incident response and recovery playbooks for emerging kinds of attacks can further improve response across organisation borders and hence has the potential to diminish the damage caused by new attacks. Additionally, shared playbooks can provide a good starting point for the automation of respective response processes.

Sharing playbooks, however, requires a mutual understanding of what a playbook is. While the common understanding is that a playbook describes a conditional sequence of steps to take for the mitigation (and sometimes also prevention or analysis) of a certain kind of incident, the specific format may vary significantly between different organisations, ranging from full text descriptions over structured text to machine-readable descriptions. A standardised machine-readable playbook format has the benefit of providing a clean interface, which allows the building of tools that take respective playbooks as input. This would also allow for easy integration of shared playbooks into locally deployed tools.

The SAPPAN project [L1] sets one of its goals to share incident handling information. While we were working on this goal, we came across a playbook standardisation effort organised within OASIS, called Collaborative Automated Course of Action Operations for Cyber Security (CACAO) [1]. Since this effort addresses the problem of providing a mutual understanding of what a playbook is by standardisation of the playbook representation and format, we decided to address the remaining problem of

Figure 1: Simplified structure of the MISP security playbook object.

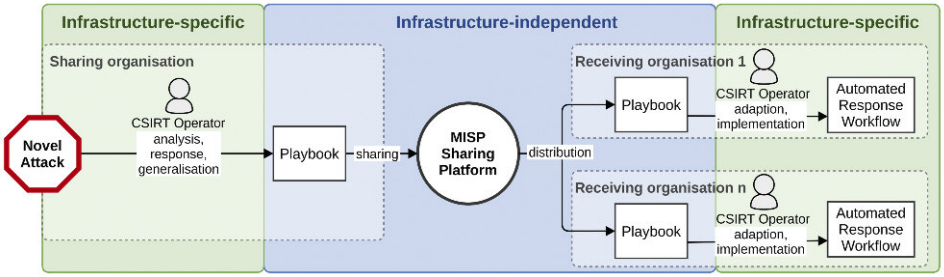


Figure 2: Overview of how collaborative playbook sharing can improve response across organisation borders. The sharing organisation detects a novel kind of attack. After mitigation, the incident is analysed and generalised to create a playbook, which is shared via MISP as a MISP security playbook object. The receiving organisations can then benefit from the sharing organisation's experience and potentially use the playbook description for preventative measures such as automated response to this kind of attack.

how to share them. We settled on Malware Incident Sharing Platform (MISP [L2]) as the sharing platform, due to its high popularity and already existing object to capture textual Course of Action (CoA) playbooks. Our goal was to prepare a MISP data model, which allows capturing standardised playbooks (such as CACAO playbooks) without being restricted to just a single standard.

In collaboration with the Technical Committee of CACAO, we prepared a MISP playbook object with specific attributes for the playbook metadata [2]. The actual standardised playbook is stored as an attachment attribute in the object as is depicted in Figure 1. This allows sharing of playbooks in other formats and does not require the transformation of a playbook when it is shared and exported from MISP. After discussing the playbook object with the MISP developers, it is now available in the official MISP object repository [L3]. We also implemented a tool to read locally stored CACAO playbooks and to correctly publish them into MISP, and vice versa.

When considering automation of incident response workflows based on shared playbooks, the shared playbooks themselves should not be specific to the sharing organisation's infrastructure. One reason for this is that if a playbook contains infrastructure-specific information, it may pose a security risk, if the receivers are not fully trusted. A second reason is that infrastructure-specific playbooks have little to no value for organisations with differing infrastructure. Playbooks suitable for sharing should hence be infrastructure-independent. Automation of incident response workflows, on the other hand, is infrastructure-specific. Consequently, there is a gap between the level of detail in which shared materials should be provided and the level of detail required for automation of respective workflows. Even though some information for incident response automation can be automatically transferred from the playbook to the automation engine (e.g., the general structure of the workflow), there is still manual work required by human operators, due to this gap in the respective levels of detail. A schematic overview of how the sharing process aligns with the automation process is shown in Figure 2.

The process of adapting a playbook to a specific organisation was carried out as a prototype for the example of responding to suspected, outgoing malware communication. To this end, the playbook was adapted to an automated workflow using Apache Airflow as a workflow engine. This automated workflow was then integrated with the case management solution used by Dreamlab's Cyber Security Operation Centre (CySOC) [L4]. The goal of the workflow is to automatically resolve as many detections as possible by blocking suspected malicious traffic from and to affected hosts, whilst simultaneously keeping the risk of disruptions low. It does this by distinguishing between critical and uncritical assets (which is done through integration with an organisation-specific asset inventory) and applying a heuristic to automatically resolve alerts which affect normal (as in not critical) hosts. This achieves two things:

1. The majority of alerts are resolved quickly and automatically with minimal risk of disruptions, which frees up analyst-time in the Security Operation Centre (SOC).
2. Alerts affecting critical hosts can be examined more closely with the now available, additional analyst-resources.

While the implementation and integration of the automated workflow into a real-world SOC environment showed that it requires non-negligible initial effort, it also revealed that it can improve response times for ordinary incidents significantly. Focusing automation efforts on incidents involving normal assets seems to provide the best trade-off between rapid incident response and the risk of disruption.

This work was done within the EU H2020 project SAPPAN [L1]. SAPPAN has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 833418.

Links:

[L1] https://sappan-project.eu/
[L2] https://www.misp-project.org
[L3] https://kwz.me/hfE
[L4] https://kwz.me/hfK

References:

[1] B. Jordan and A. Thomson eds.: "CACAO Security Playbooks Version 1.1" OASIS Committee Specification, 22 Oct. 2021. Available online: https://kwz.me/hfH
[2] V. Mavroeidis, et al.: "On the Integration of Course of Action Playbooks into Shareable Cyber Threat Intelligence", IEEE Big Data, 2021.

Please contact:

Lasse Nitz
Fraunhofer Institute for Applied Information Technology (FIT), Germany
lasse.nitz@fit.fraunhofer.de

Towards Model-Driven DevSecOps for Cyberattack Prevention, Detection and Recovery

by Christophe Ponsard, Philippe Massonet, Valery Ramon (CETIC)

Coping with cybercrime in the scope of increasingly open and interconnected systems is a difficult challenge. DevSecOps provide an adequate framework to keep in control of this perpetual race. We show here how it can be efficiently supported by an internal model-based analysis and automation approach together with the external threat intelligence sharing.

An effective way to counter cyberattacks targeting a company's assets is to organise several lines of defence based on a strong risk culture. Typical lines are to protect, detect/respond, and be able to recover, as documented in the NIST Cyber Security Framework [L1]. However, in order to be efficient, organisations need to be as reactive as the cybercriminals, who have structured themselves along a whole value chain including selling vulnerability exploits, providing attacks as a service, and making profit based on stolen or ransomed data.

The DevSecOps approach aims at creating and including modern security practices in the fast and agile world of DevOps [1]. It promotes collaboration between developers and operators by involving security experts, while relying on strong automation (e.g. bots) and continuous integration/delivery tools. To make it even more efficient, our recent research explored the use of modelling techniques in order to build consistent and integrated model chains [2] as depicted in Figure 1.

Key modelling activities start from threat modelling at the planning phase and are managed first throughout the development phase, including making sure the code is immune to common attack patterns (CAPEC) and common weaknesses (CWE), before going into the build and test phase using building automation and targeted pen testing. At this level, in the scope of the SPARTA project [L2], we are also developing models and tools to support incremental certification processes, which provide third-party assurance on the security of the deployed solution [3]. Then, in the operations phase, an infrastructure model, similar to the planning phase, is used as well as access control models for secure operations. Active monitoring is also used to detect known or unknown attack types, possibly through shared Common Vulnerability and Exposures (CVE), and to trigger automated responses to them.

The knowledge acquired through the monitoring, especially unknown attacks, can be reported by the local Secure Operation Centre (SOC) to a network of Computer Emergency Response Team, which can result in new CVE, which can lead to documenting new CWE and CAPEC. This learning part occurs outside the organisation and proceeds backwards through the development phases, enabling the coping with new security issues as early as possible.

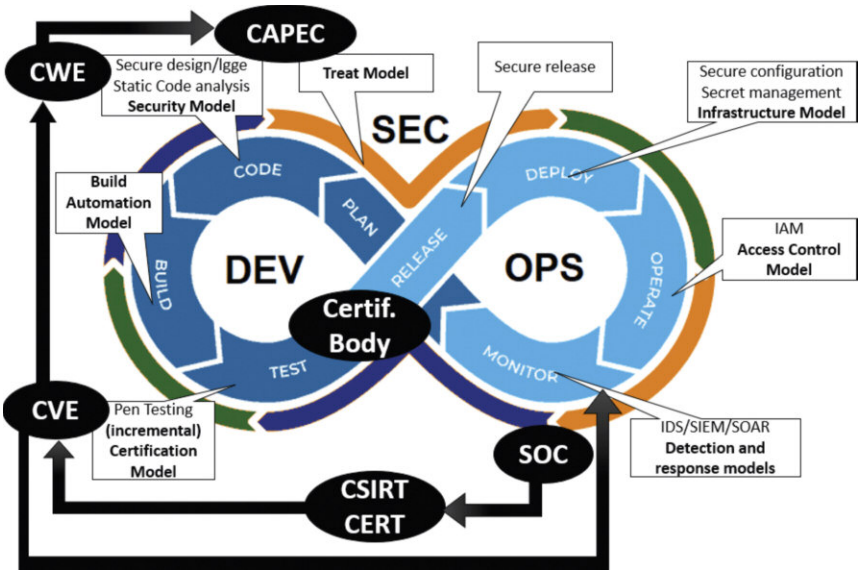


Figure 1: DevSecOps model-oriented activities in internal and external chains.

The next steps of our research are based on grand challenges proposed by CyberWal platform, driven by industrial needs [L3]. They target problems such as the precise risk modelling of Cyber Physical Systems, automation of attack scenarios for driving penetration testing, and AI-based intrusion detection for industrial systems.

Links:

[L1] https://kwz.me/hfD
[L2] https://www.sparta.eu
[L3] https://cyberwal.be

References:

[1] H. Myrbakken, R. Colomo-Palacios: "DevSecOps: a multivocal literature review", SPICE, 2017.
[2] J. Hugues, J. Yankel: "From Model-Based Systems and Software Engineering to ModDevOps", CMU, 2021.
[3] D. Sébastien, et al.: "Incremental Common Criteria Certification Processes using DevSecOps Practices", EuroS&P Workshops 2021.

Please contact:

Christophe Ponsard
CETIC, Belgium
+32 472 56 90 99
christophe.ponsard@cetic.be

Strengthening the Mobile Forensics Investigation Chain

by Phil Cobley (MSAB), Georgina Humphries (NMPS), Harry Manifavas (FORTH-ICS), Rune Nordvik (NMPS), Matthew Sorell (Univ. of Adelaide)

Mobile devices, especially smartphones, constitute a major source of evidence in criminal activities investigated by law enforcement agencies (LEAs) [1]. Mobile devices present unique challenges; therefore, it is vital to empower all players involved in solving and judging cases where mobile data plays a significant role. FORMOBILE [L1], an EU-funded H2020 project aims to establish a complete end-to-end forensic investigation chain for mobile devices by developing the first standard for mobile forensics, novel tools, and a targeted training programme.

The wider population increasingly utilises smartphones in their daily routine. Criminals also use mobile devices to coordinate their illegal activities. According to recent statistics [L1], 85% of crime investigations include mobile data.

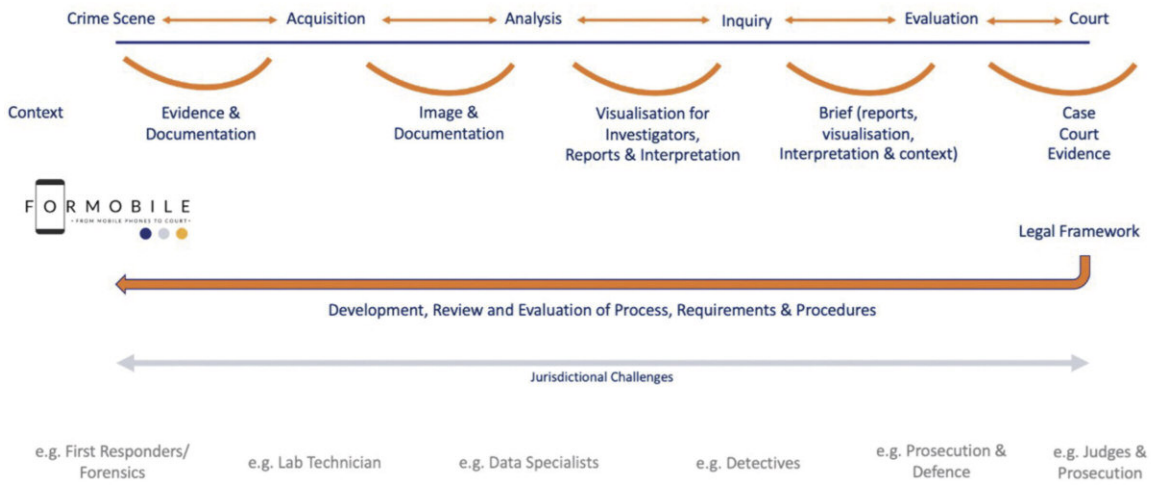
Digital forensic investigators respond to a crime committed in the physical or cyber space by trying to reconstruct the crime based on the content stored on the digital devices seized and to produce sound evidence that can stand scrutiny in court. The full (end-to-end) investigation chain is composed of several phases and involves different types of practitioners (Figure 1). The phases can be categorised as Crime Scene, Acquisition, Analysis, Inquiry, Evaluation and Court. Practitioners include First Responders, Lab Technicians, Data Specialists, Detectives/Investigators, Prosecution, Judges and Defence.

The FORMOBILE standard "Requirements and Guidelines for a complete end-to-end mobile forensic investigation chain" fills a gap in the standardisation activities around Digital Forensics. It aims to codify good practices observed in the field of mobile forensics. Practitioners adhering to the standard demonstrate a consistent and justified approach to dealing with mobile data. This is an essential requirement for the data to be considered as admissible evidence before the relevant court.

FORMOBILE tools allow LEAs to more robustly retrieve data stored on mobile phones, better decode retrieved data (especially types of data previously off-limits), and finally more efficiently visualise and analyse decoded data. Challenging data sources are for example, devices employing encryption or anti-forensics measures, cloned phones, and cloud accounts.

FORMOBILE, together with LEAs, commercial entities, and academia, developed a novel mobile forensics training curriculum. The curriculum defines a recommended set of courses for an individual to become a practitioner in mobile forensics. During the project, partners – led by the Norwegian Police University College on behalf of the Norwegian Ministry of justice and Public Safety (NMPS) – developed several online e-learning courses. Each course consists of lessons grouped into modules. The courses target all the practitioners listed in Fig. 1. More specifically, the courses developed are: Mobile Forensics Fundamentals and Best Practices, Mobile Forensics for Management, Mobile Forensics using FORMOBILE Tools. Current activities also include the development of two more courses, Mobile Forensics for Prosecution and Judges and Mobile Networks. Proof-of-concept trainings with LEAs and a train-the-trainer week-long event have been performed to help disseminate knowledge and practice, evaluate the trainings, and receive feedback.

The pilot training concluded with a week-long Capture the Flag (CTF) competition where the participants put their knowledge and skills into practice. During the CTF, participants were presented with a case scenario and a set of tasks consisting of several challenges. To answer each question, each participant had to reflect on the background theory



and use the appropriate tools to extract relevant evidence from the device extractions provided. The CTF setup offers a gamified learning experience.

The overall aim of the training is to increase the quality of investigations by making the targeted audience aware of the complexities that surround such investigations, standardisation efforts, forensic tool intricacies, the evidential value of technical artefacts as well as the applicable legal frameworks and provisions (e.g., principle of necessity, principle of minimisation, right to a fair trial).

The FORMOBILE consortium consists of nineteen partners (L1). It brings together LEAs, commercial companies, civil organisations, and academic institutes. The project partners come from thirteen EU countries and two associated countries. The project started in May 2019 and ends in April 2022. Some of the results from the training will be made publicly available at the end of the project for the mobile forensic community, while other results will only be available to LEAs for reasons relating to security.

Link:
[L1] https://formobile-project.eu/

Reference:
[1] Dismantling of an encrypted network sends shockwaves through organised crime groups across Europe, Aug. 2020. https://kwz.me/hfI

Please contact:
Harry Manifavas
FORTH-ICS, Greece
harryman@ics.forth.gr
communication@formobile-project.eu

Identifying Attack Propagation Threads and Root Cause in Internet-of-Vehicle Ecosystems Utilising Honeypots in Connected Autonomous Vehicles

by Christos Alexakos (ISI/ATHENA RC), Kristina Livitckaia (ITI/CERTH), Mike Anastasiadis (ITI/CERTH), Dimitrios Serpanos (University of Patras)

The importance of cybersecurity for Internet of Vehicles (IoV) systems is indisputable as possible attacks can cause the loss of lives. In nIoVe project, a cybersecurity framework has been developed. This framework includes tools for accurate detection of the propagation trends and root cause analysis of the attacks, providing additional knowledge for cyberattacks against lookalike infrastructures.

Autonomous vehicles mainly operate inside a smart city infrastructure. Due to the plethora of Internet-connected entities inside and outside the vehicle, the cybersecurity issues emerging are of utmost importance. In cybersecurity, the knowledge of previous attacks strengthens the systems that are responsible for the detection and identification of a cyberattack as well as the systems orchestrating the mitigation actions. This article presents an approach that utilises attack data collected by honeypots installed inside the vehicle to define the root cause and the propagation trends of detected attacks. The scope is to define, through a harmless infrastructure, a honeyfarm, the possible root cause of a cyberattack on an autonomous vehicle and share this information with other detection tools for faster and more reliable intrusion detection. The presented tools are part of the security infrastructure of nIoVe: A Novel Adaptive Cybersecurity Framework for the Internet-of-Vehicles European project [L1] by researchers of Industrial Systems Institute of ATHENA RC and Information Technologies Institute of CERTH, Greece.

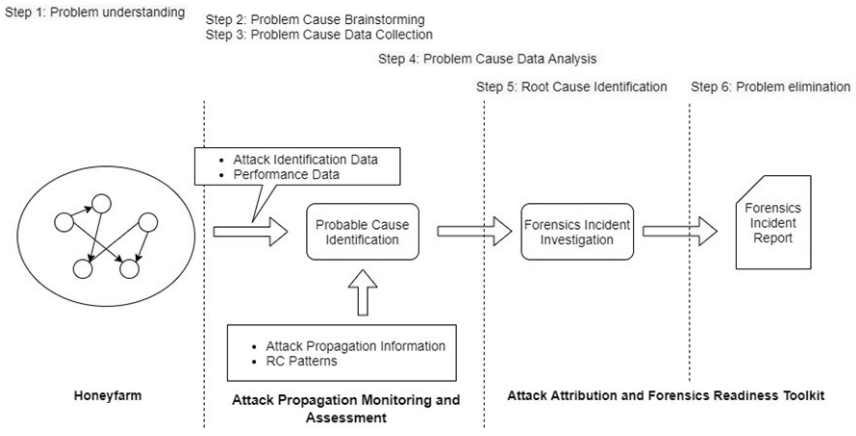


Figure 1: The Attack Propagation Monitoring and Root Cause Analysis Procedure of the nIoVe framework.

The approach that was followed is depicted in Figure 1. The approach follows the steps that define the Root Cause Analysis (RCA). According to Wangen et al. [1], RCA is considered a structured investigation that follows a well-defined procedure to identify the cause of the detected problem and define the actions to eliminate or prevent it in the future. Andersen and Fagerhaug [2] presented a simplified vision of the root cause problem where the identification of the root cause is a sequential process looking at the bottom reason that initiated the cause-and-effect chain that finally causes the problem.

The first phase is the installation of the honeypots' infrastructure. A tool has been developed to allow administrators to create a honeyfarm dynamically. The honeyfarm will be a combination of different honeypots emulating the sensors

inside the vehicle and generally the autonomous vehicle entity (Lidar, camera, CAB bus, IMU, OBU, etc.). The utility of the honeyfarm is to attract the attackers into attacking it while allowing the monitoring of the attack mechanisms used against it.

The most common way to detect propagation trends is to apply Markov Chain Models. A Markov Chain is a stochastic process that assumes the next step of the process depends only on its current state. In the context of detecting propagation trends, Markov Chains are applied to model attack sessions, with each node representing a honeypot being attacked and the transitions showing that another honeypot is attacked immediately after the last. The following process is adopted to build a Markov Chain Model based on the data collected by a honeypot framework: a) Single events detected in the framework are used to construct attack sessions. b) The attack sessions are used to construct the Markov Chain Model. Also, contextual models can be built by filtering the list of single events to only include events happening within a certain context. Once the Markov Chain Model graph has been constructed, graph analysis techniques based on clustering algorithms and centrality metrics are applied to gather insight into the system's propagation trends.

The procedure of RCA is semi-automated and based on six steps as depicted in Figure 1. The pipeline starts when an anomaly behaviour is detected on the honeyfarm. Based on the previous knowledge of the attack propagation and with the data about the performance of the systems in the network, the process will identify the probable honeypots and vulnerabilities that are the root cause of the attack. This approach uses a semantic graph that depicts the networked system of the honeyfarm. Each node in the graph is a system/device, each edge provides information about the relations between two nodes, each object's properties define the functionalities of the nodes. For the first step in the RCA, the procedure is Problem Understanding. For this, a Performance vs. Importance matrix will be created. The matrix will be a 2-dimensional diagram. The x-axis will represent the values of each node's functionality performance, and the y-axis the importance of the node's functionalities. For the RCA, the number of the points in the quarter Bad Performance and High Importance defines the problem and identifies the probable nodes that cause it. The next phase of RCA is a data collection analysis process, including the collection of the attack data and its process in two steps. First, the use of the attack propagation knowledge to back trace the propagation pattern to the probable root cause. Second, the use of graph similarity methods to find information from previous attacks to similar systems. These algorithms aim to detect the node and the functionality (e.g., network, software services, etc.) that cause the problem to the overall system. Finally, the data collected and the identified causes will be forwarded to the Forensics Assessment tool where a security expert will investigate all the information and generate a report including the RCA and suggestions for mitigation actions, which are finally saved in a shared threat intelligent repository based on MISP platform [L2].

The presented approach can be used to model valuable information of possible attacks on a complex and sensitive system such an autonomous vehicle. The research on the collection of data from various attacks and the optimisation of the attack propagation trends and root cause is a continuous challenge for the authors, who anticipate automating the semi-automatic process of RCA.

Links:

[L1] https://www.niove.eu

[L2] https://www.misp-project.org/

References:

[1] G. Wangen, et al: "An empirical study of root-cause analysis in information security management", SECURWARE, IARIA pp. 1-14, 2017.

[2] B. Andersen and T. Fagerhaug: "Root cause analysis: simplified tools and techniques", Quality Press, 2006.

Please contact:

Christos Alexakos, Industrial Systems Institute/ATHENA RC, Greece
alexakos@isi.gr

PenQuest: Gamifying Cyberattacks

by Robert Luh and Sebastian Eresheim (University of Vienna & St. Pölten University of Applied Sciences)

PenQuest is a digital multi-player game that allows users to recreate or emulate cyberattacks on a game board representing freely configurable IT infrastructures. The game's model incorporates a multitude of security concepts and threat vocabularies translated into technical and organisational actions. PenQuest is intended to assist risk assessment, support the reconstruction of adversarial events, and gamify security education.

PenQuest [1,2] is an adversarial cybersecurity game conceived and developed by St. Pölten University of Applied Sciences and the University of Vienna, with additional funding from DIH-OST. It combines a board game modelling freely configurable, hierarchical IT infrastructures with game actions representing technical, organisational, and social engineering attacks as well as appropriate defensive measures. PenQuest was originally built to support security education and awareness trainings but has long since branched out to encompass risk assessment and forensic event reconstruction.

Under the hood, the attacker's potential actions are derived from the MITRE ATT&CK framework [L1], while defensive actions are based on MITRE D3FEND [L2] as well as the NIST SP 800-53 security standard [L3]. The effects of data theft (confidentiality attacks), system manipulation (integrity

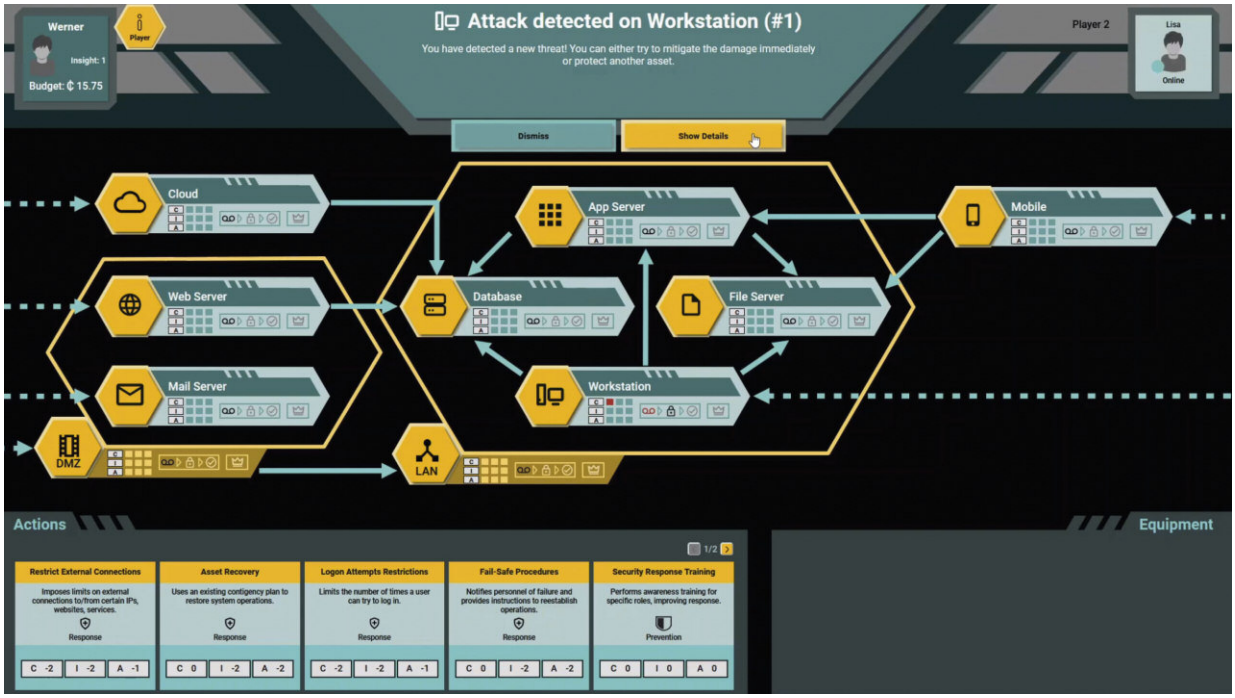


Figure 1: PenQuest game board depicting a common network infrastructure.

attacks), and attacks on availability are modelled in addition to the different kill chain phases of an attack, which typically range from reconnaissance, to propagation, and "detonation" activity. Simply put, actions deal a certain amount of "damage" and are limited to one or several attack stages in accordance with their technical specifics. For example, the action "Inter-Process Communication" that abuses data exchange mechanisms between processes to execute code can only be used once the attacker has gained access to the system in an earlier turn.

Attacking and defending actors of various levels can be defined using numeric attributes denoting skill, motivation, and financial means. These attributes in turn limit available actions and equipment (security appliances, attack tools, etc.) that can be utilised or procured. Furthermore, the game considers the inter-dependencies of systems, privilege levels, and models vulnerabilities that can be exploited through e.g., malicious software procurable in-game.

Another key part of the framework revolves around the mapping of mitigating controls to specific attacks in order to maximise realism and help design more effective prevention, detection, and response measures. To this end, various approaches ranging from abstracting attacks to their base components (e.g., account elevation, scanning, configuration changes, or social engineering) to natural language processing (NLP) were explored and evaluated with the help of security experts.

Put together, a multitude of real-world threat scenarios can be depicted within the context of PenQuest and combined to reconstruct the events leading up to a system compromise, to create awareness-raising measures, or even to conduct comprehensive risk analyses. Thanks to the unrestricted configuration of actors, actions, and systems, there are hardly any limitations – from modelling a ransomware attack on an isolated workstation performed by an opportunist individual, to recreating a large-scale data theft by a well-funded corporate competitor, anything is possible.

While the PenQuest server and client are already available for use to early adopters and testers (see figure), ongoing research conducted as part of the "INODES" project [L4] at University of Vienna focuses on optimising strategies through reinforcement learning and model checking. We specifically aim to determine which attacks are most likely to succeed given different financial, temporal, or skill constraints and how defensive measures can best protect against them. A tie-in of the model to existing intrusion detection systems (IDS) is in the works as well. This way, it will become possible to link current alerts or indicators of compromise identified during a forensic investigation to likely attacker motivations and objectives, enabling not only the interpretation of past cyber-threats, but also the timely response to ongoing attacks.

Visit [L5] for more information about PenQuest and contact the article authors for early access to the game.

Links:
[L1] https://attack.mitre.org
[L2] https://d3fend.mitre.org
[L3] https://kwz.me/hfS
[L4] https://kwz.me/hf4
[L5] https://www.pen.quest

References:
[1] R. Luh et al.: "PenQuest Reloaded: A Digital Cyber Defense Game for Technical Education", in Proc. of EDUCON Conference, 2022.
[2] R. Luh et al.: "PenQuest: a gamified attacker/defender meta model for cyber security assessment and education", Journal of Computer Virology and Hacking Techniques, 2019.

Please contact:
Robert Luh
University of Vienna and St. Pölten University of Applied Sciences, Austria
robert.luh@univie.ac.at

Sebastian Eresheim
University of Vienna and St. Pölten University of Applied Sciences, Austria
sebastian.eresheim@univie.ac.at

Fighting Cybercrime through Education: Integration of an Educational Cyber Defence Centre into Cyber Security Curricula

by Jochen Hense, Simon Tjoa, Peter Kieseberg (St. Pölten University of Applied Sciences, Austria)

Traditional security education often focuses on teaching theoretical concepts but lacks hands-on experience. However, many aspects of modern security, especially in the incident management domain, cannot be taught in the abstract; they must be experienced. Our Cyber Defence Centre (CDC) allows us to train students in a simulated environment where they gain skills in detecting attacks, closing vulnerabilities, and responding to security breaches in a realistic but safe setting.

Modern society and economies are highly dependent on secure, reliable and available infrastructures. Cybercrime has evolved into a serious threat, ranging from primitive "spam & scam" to sophisticated targeted attacks that focus on a specific target or an entire industry. This is particularly facilitated by the global and autonomous nature of the Internet, which makes geographic distance meaningless in the event of an attack. Critical infrastructures therefore require additional protection, as downtime in these systems affect the lives of many citizens. Moreover, not only direct attacks on such an infrastructure need to be carefully considered, but the distributed infrastructure can also be used as a platform for further attacks on other industries. Thus, the European Union has established the EU NIS- directive [1], aiming at increasing the resilience of infrastructure deemed critical and facilitating the exchange of vital security-related information. This European dimension also means that infrastructure protection is becoming an interconnected problem, involving multiple players in a field as well as cross-sector supply chains.

Cyber Defence Centres (CDCs) [2] play a central role in the protection of information systems, as they integrate information from inside the enterprise as well as from external sources in order to establish situational awareness of the current risk situation. CDCs offer services for the structured collection, processing and provision of security-critical information to decision-makers and analysts. Using the integration of external information, machine learning and other AI techniques for pattern detection, it is often possible to detect attacks early (see Figure 1). Thus, especially large enterprises and government organisations consider cyber defence centres as the cornerstones of their coordinated defence activities.

In particular, the so-called "war rooms" [3], the central points where decisions are made on the basis of visualised information aggregates, are familiar from many presentations by relevant providers. However, many important analysis steps take place separately by trained personnel: In this context, the use of AI for (semi-) automation of information analysis is becoming more and more important.

For this reason, there is high demand for experts that not only possess technical provenience, but are also familiar with cyber defence operations. The major problem is that many important aspects of real-life cyber security, especially related to stress and decision-making based on limited information, cannot be taught in the classroom, but need to be experienced in a setting that is as close to actual real-life situations as possible. Still, CDCs are currently hardly present in traditional tertiary education in the field of security. Setting up an educational CDC is far from trivial, as there are several challenges to be solved:

- The infrastructure required to simulate real life requires a dedicated infrastructure.
- For the simulation of incident response, a so-called "war room", a dedicated room for information visualisation and decision-making, has to be established.
- In addition, dedicated workplaces for analysts need to be set up that allow for the analysis of complex attack scenarios in the midst of a sea of information and sensor data.
- Since in an educational setting, specific attack scenarios against well-defined infrastructures need to be taught and analysed, these attacks and infrastructures need to be simulated. Thus, a simulation engine allowing for varying scenarios needs to be developed.
- In addition, simulations purely based on artificially conceived data lack details due to the enormous amount of data that is available in real life

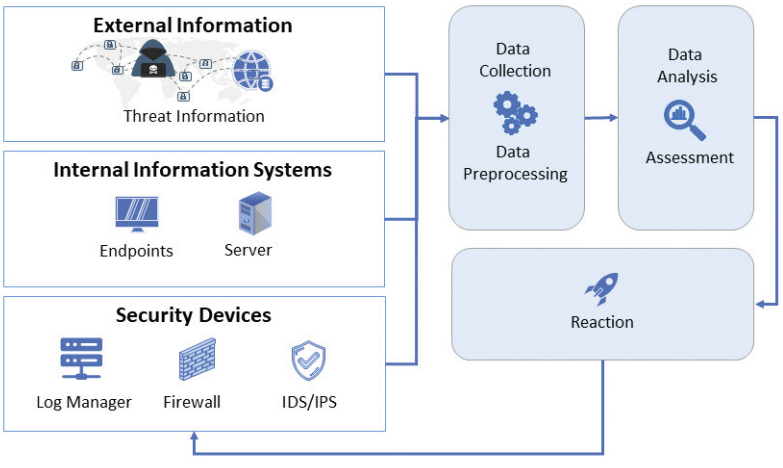


Figure 1: Integration of external information and AI into the CDC.

attacks. Thus, it is beneficial to use data gathered through actual attacks in the simulations that needs to be packed in the form of well-defined scenarios fit for training the students.

The teaching CDC works like a "teaching hospital" with the approach of "clinical clerkship" and "residency". In this approach, students will learn on the job with pseudo-real data and incidents under the guidance of the teaching faculty. This puts students in a position, upon graduation, to strengthen the cyber security field, which now has much more demand than there are highly qualified professionals available.

This is exactly where the “Educational CDC” (eduCDC) of the FH St. Pölten comes into play. By integrating the teaching CDC into a wide variety of courses on a wide variety of topics and with the help of different scenarios, our students can gain this important experience, and they can virtually move in a (simulated) real situation, with all the additional characteristics available, such as stress, unclear objectives and technical challenges. This also provides much-needed practical experience (hands on the job) as an optimal preparation for professional entry, with a focus on the following key competencies of our students: systems monitoring in order to detect vulnerabilities and ongoing attacks, digital forensics of system components to find the root causes of security incidents, penetration testing for the proactive detection of vulnerabilities, malware analysis for the detection of (previously unknown) malicious software ,and threat intelligence that analyses state-of-the-art attacks in order to be able to react in a timely manner and with profound countermeasures. Aside from this technical level, we also focus on security management for the identification and assessment of risks, as well as disaster recovery, focusing on the restoration of systems after successful attacks. This is all rounded off with lessons in training security awareness for reducing the human attack surface.

Summarised, the introduction of the eduCDC into our curricula is a novel concept enabling us to provide our students with the most cutting-edge security education.

Link:
[L1] https://kwz.me/hfR

References:

[1] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union
[2] Miller, J.N. and Butler, R.J., 2021. National Cyber Defense Center: A Key Next Step toward a Whole-of-Nation Approach to Cybersecurity. Johns Hopkins University-Applied Physics Laboratory.
[3] O'Reilly, C., Bustard, D. and Morrow, P., 2005, May. The war room command console: shared visualizations for inclusive team coordination. In Proceedings of the 2005 ACM symposium on Software visualization (pp. 57-65).

Please contact:

Peter Kieseberg
St. Pölten UAS, Austria
peter.kieseberg@fhstp.ac.at

Ludoteca del Registro.it: Cybersecurity in Education

by Giorgia Bassi, Stefania Fabbri and Anna Vaccarelli (IIT-CNR)

Ludoteca del Registro.it is a project implemented by the Registro.it (the Registry of .it Internet domains) of the Institute of Informatics and Telematics of the CNR (National Research Council) in Pisa, aimed to help students develop more responsible use of the Internet, with a focus on cybersecurity topics.

Ludoteca del Registro.it [L1] [L2], sponsored by the Garante per l’Infanzia e l’Adolescenza (the Italian Authority for Children and Adolescents) is a project to promote Internet culture. So far, the project has engaged about 500 classes and 14,000 students throughout the country, during educational workshops. It is aimed at schools from primary through to high schools, but teachers and parents can also find information and resources on the website. The purpose is to give a complete view of the Internet and understand what it is, why to use it and how to use it properly.

It is, therefore, crucial to make users, and especially the youngest ones, aware of the potential risks and encourage them to adopt behaviours that can avoid them. After all, human behaviour is still the weakest link in cybersecurity, as demonstrated by the rise of social engineering attacks such as phishing.

Providing the basis of Digital Culture and Cybersecurity skills, not only related to technical aspects but also extended to cultural ones, is crucial to create awareness and, therefore, to prevent cyber threats and risks.

The educational workshops are mostly focused on cybersecurity, but they also explain the technical infrastructure (TCP/IP protocol, IP address, packet switching, internet domain names, DNS system), in order to better master its use and take advantage of its opportunities.

Ludoteca has different means for different levels of school: the web app Internetopoli (the city of Internet), for primary school; the videogame "Nabbovaldo e il ricatto dal Cyberspazio" (Nabbovaldo and the blackmail from Cyberspace) and the related educational path for secondary schools; and Cybersecurity4Teens, comprising 10 hours of lessons and practice about Cybersecurity for high school. Finally, there is the training web portal "Presente Digitale" addressed to teachers, dealing with some digital top-

ics, with the aim to help them master these arguments. All the initiatives are free.

The Internetopoli web app

One of the main tools of the project is Internetopoli, a multimedia application about the Internet, which is freeware and suitable for interactive whiteboards [L3] [1].

Figure 1: Internetopoli.

Internetopoli is the Internet city, a metaphor that links together all the macro themes of the project. The main topics, explained though the metaphor of the city, are listed below:

- How the Internet works (Internetopoli is a city composed of streets, houses, and addresses, just as the Internet is made up of computers and devices, linked through special addresses called IPs)
- The domain names (addresses of the Internetopoli houses are the internet domain names, easy for the citizens to memorise)
- Governance and organisation of the Internet (Internetopoli, like all cities, is regulated and administered by local and international institutions)
- Safe and knowledgeable use (the citizens of Internetopoli must be responsible and respect rules in order to protect their privacy and that of others)
- Opportunities of the Internet (Internetopoli is a city rich in resources and services, useful for everyday life, work and study)

One of the most important concepts in the App is the “Internet citizenship”, to be interpreted as a community totally involved in a safer and responsible use of the Internet.

Cybersecurity4Teens

The "CyberSecurity for Teens" (CS4TY - [L4]) is a framework involving the 11–19 age group (secondary and high school students) with the purpose of putting together a series of resources and activities dedicated to cybersecurity issues. The goal is to acquire a vertical curriculum dedicated to cybersecurity, focusing on the following skills: protecting devices; protecting personal data and privacy; recognising and fighting the risks of "cyberspace", e.g., being aware of the interaction of people, software and services by means of technologies, devices and networks connected to it.

These skills will be achieved through the acquisition of knowledge of: systems threats, vulnerabilities and cyberattacks; and technical countermeasures such as authentication, cryptography and identification and intrusion prevention systems (antivirus and firewall).

The importance of the acquisition of such skills has led to the conception of CS4T in such a way that it will be extremely customisable according to the age groups involved, integrating different resources and methodologies.

The secondary school labs are based on the use of the videogame "Nabbovaldo and blackmail from cyberspace", a tool to introduce teenagers to cybersecurity topics in a funny and engaging way (see the following section).

High school labs are developed in three webinars (the first is about the main cyber threats, the second about the countermeasures, and the third about the main types of hackers) and a workshop held by some researchers of the Trust, Security and Privacy Unit of IIT-CNR [L5], with a total of ten hours.

The videogame: “Nabbovaldo e il ricatto dal Cyberspazio”

The videogame “Nabbovaldo and the blackmail from Cyberspace” [L6] has been designed for secondary schools [2]. The Nabbovaldo character was created for two comic books (“Nabbovaldo, ovvero le stagioni a Internetopoli” – “Nabbovaldo and the seasons in Internetopolis” and “Nabbovaldo contro i pc zombi” – “Nabbovaldo versus the zombie laptops”) written and illustrated by Gabriele Peddes. Nabbovaldo moves in

Figure 2: The videogame "Nabbovaldo e il ricatto dal cyberspazio".

Internetopolis, seeking his fortune; he is an inexperienced guy who must face the challenges and the opportunities offered by an unknown metropolis. In particular, he faces a ransomware that infected all the city. The videogame (a single player) deals with many cybersecurity topics: malware, hate speech, fake news, trolls, dark web, sexting, etc. The videogame is an opportunity for teachers to explore these topics with pupils and explain countermeasures.

Teachers

The Ludoteca del Registro.it project includes the web portal “Presente Digitale” [L7], that hosts some online training courses, addressed to teachers. The courses are "Teaching by means of and in the net", "Computational thinking and Coding", and "Cybersecurity".

Links:

[L1] https://www.ludotecaregistro.it/
[L2] https://www.nic.it/en
[L3] https://www.internetopoli.it/
[L4] https://kwz.me/hjO
[L5] https://tsp.iit.cnr.it/en/
[L6] https://kwz.me/hjk
[L7] https://presentedigitale.it/

References:

[1] R.M. Bottino, et al.: “Digital games in primary schools for the development of key transversal skills”, in Proc. of SUZA, 2019.
[2] L.S. Ferro, et al: “A game-based learning experience for improving cybersecurity awareness”, 4th Italian Conference on cybersecurity, Itasec 2020, Ancona, Italy, 4-7 Feb 2020, pp. 235-242.

Please contact:

Giorgia Bassi
IIT-CNR, Italy
giorgia.bassi@iit.cnr.it

European Research and Innovation

Artificial Intelligence Enabled Distributed Edge Computing for Internet of Things

by Ali Balador, Sima Sinaei (RISE Research Institute of Sweden) and Mats Pettersson (Sensative AB)

DAIS is a huge step forward in the area of artificial intelligence and edge computing. DAIS intends to create a complete framework for self-organising, energy-efficient and private-by-design distributed AI. DAIS is a European project with a consortium of 47 partners from 11 countries coordinated by Ali Balador from RISE research institute of Sweden.

In recent years, technological developments in consumer electronics and industrial applications have been advancing rapidly. More and smaller, networked devices are able to collect and process data anywhere. This Internet of Things (IoT) is a revolutionary change for many sectors like building, automotive, digital industry, energy, healthcare, etc. As a result, the amount of data being generated at the edge level has and will increase dramatically, resulting in higher network bandwidth requirements. In the meantime, with the emergence of novel applications, such as automated driving, lower latency of the network is required.

The new paradigm of edge computing (EC) provides new solutions by bringing resources closer to the user, keeps sensitive & private data on device, and provides low latency, energy efficiency, and scalability compared to cloud services, while reducing the network bandwidth, in addition bringing cost savings. EC guarantees the quality of service when dealing with a massive amount of data for cloud computing [1,2].

At the same time, Artificial Intelligence (AI) applications based on machine learning (especially deep learning algorithms) are being fuelled by advances in models, processing power, and big data. In Cisco's annual report (2018-2023) [L1], all 83 organisations asked reported that they have edge computing use cases where artificial intelligence, IoT and 5G had higher portions. The huge increase of devices at the network edge drives the need for enterprises to manage and analyse data from IoT endpoints. Shifting traffic from the network core to the edge affects computing and communications architectures. To have a successful edge computing strategy, it is important to make sure the overall infrastructure is efficient and manageable.

The development of AI applications mostly requires processing of data in centralised cloud locations and hence cannot be used for applications where milliseconds matter, or for safety-critical applications. For example, as the sensors and cameras mounted on an autonomous vehicle generate about a gigabyte of data per second, it is difficult, if not impossible to upload this data and get instructions from the cloud in real-time. The same applies to face recognition applications

The DAIS consortium.

that have high temporal requirements for processing either online or offline. Moreover, edge computing offers security benefits due to wider data distribution at the edge level. Reducing the distance data has to travel for processing means decreasing the opportunities for trackers and hackers to intercept it during transmission and preserving its privacy. With more data remaining at the edges of the network, central servers are also less likely to become targets for cyberattacks. This has led to a growing interest in Federated Learning (FL), as a promising distributed learning paradigm that allows multiple parties to jointly train a global ML model on their combined data, without any participants having to reveal their data to a centralised server [3].

DAIS is a huge step forward in the area of artificial intelligence and edge computing. DAIS aims at providing edge computing architecture, including both hardware and software, for industrial applications. DAIS is a pan-European effort spanning three years with a total budget of €33 million and a consortium of 47 partners from 11 countries. Coordinated by Ali Balador from RISE, Research Institute of Sweden, and with the support of Europe's industry, Europe's leading research organisations, the European Union via the KDT Joint Undertaking and the participating national funding agencies, it is possible to bring together European and International key players to the benefit of Europe's economy and society.

DAIS intends to create a complete framework for self-organising, energy efficient and private-by-design distributed AI. The framework covers an end-to-end system consisting of a variety of heterogeneous nodes ranging from simple IoT nodes to high-performance servers and cluster nodes. The framework consists of a hardware framework and a software framework enabling both traditional end-to-end services as well as horizontal systems with multiple service providers using common data, IoT and AI resources. Multiple connected nodes using the DAIS framework creates a complete DAIS system supporting both vertically and horizontally distributed AI. The DAIS framework will also support open standards to interact with other devices, platforms, and services. In DAIS, we leverage existing hardware devices and refine them, as well as develop new devices.

The DAIS SW framework includes a DAIS AI framework, a DAIS security framework, and a DAIS communication framework including SW components, tools, policies, interfaces, and data-models enabling compatibility, privacy, and secure data communication between products and services, creating a fully distributed AI system. The DAIS framework will allow for standardised data pipelines, sharing compute load between nodes, model distribution, as well as model learning hierarchies such as federated learning. Furthermore, it will provide a standardised way of sharing, keeping track of, and distributing model-specific configurations, data and model architectures in a secure manner. The framework is producer-, generator-, vendor- and service-agnostic, hence allowing interaction between the different nodes vertically and horizontally.

Link:

[L1] https://kwz.me/hfU

References:

[1] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp.637–646, 2016.

[2] Ding, Aaron Yi, et al. "Roadmap for Edge AI: A Dagstuhl Perspective." arXiv preprint arXiv:2112.00616 (2021).

[3] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50-60.

Please contact:

Ali Balador, RISE Research Institute of Sweden, Sweden
ali.balador@ri.se

Potential Hazard of Accidental Radioactive Discharges into the Calm Atmosphere

by Petr Pecha, Miroslav Kárný, Emilie Pechová, Václav Šmídl and Ondřej Tichý (Institute of Information Theory and Automation)

A research team from the department of Adaptive Systems of ÚTIA [L1] has solved the project HARP [L2] which included examination of various release scenarios under worst-case meteorological conditions. Recently, the team has been focused on inspection of the calm situations characterised by stable atmosphere at very low wind speed with possibility of rainfall. Although the probability of such episodes is low, possible radiological impact on the environment can be serious. Developed methodology supports deployment of the sampling-based methods for probabilistic estimation of the radiological impact of radiation accidents.

Accidental discharges of radioactive substances from an elevated source into the motionless (calm) atmosphere are examined with the aim to quantify possible radiological impact on the population. In a few hours of a calm meteorological situation, a significant level of radioactivity may accumulate around the source. The inspected scenario assumes that the calm situation expires and a wind-induced convective movement of the air immediately begins. Randomness of the model parameters of this CALM scenario allows generation of complex random radiological trajectories. They constitute the deposition of radionuclides on the ground during both calm and convective stages of the release [1]. Their inspection is the necessary prerequisite for the prospective uncertainty and sensitivity analyses.

Speedup based on the new "super-puff" concept

The radioactive release from an elevated source into the motionless ambient atmosphere is approximated by a long sequence of discrete discharges (instantaneous puffs) described by 3-D Gaussian-puff formulae. Each puff has a shape of an expanding discus with its centre at the pollution source. The puff distribution describes the radioactivity concentration in air released at effective height. The radioactivity depletion from the air Ḣ caused by physical mechanisms of radioactive decay, dry depletion and wet depletion Ḣ is implemented. Demands on computing resources are substantially decreased by projecting the non-Gaussian sum (red line in Fig. 1) of all partial puffs (blue lines in Fig. 1) on one representative Gaussian "super-puff" distribution. This approximation based (AB) solution uses the projection approach implied by Bayes' paradigm [2]. Benefit of the reduced computational load is evident, especially for a large number of puffs M. It requires only one-shot run modelling of the ensuing convective transport, which significantly lowers prospective demands on necessary application of more powerful but laborious dispersion codes. The brute force solution (BF), in which each puff individually undergoes the convective propagation and enters the final superposition, is obviously time consuming for a large number of discrete puffs.

Simple demonstration example

Let a certain calm episode lasts five hours and immediately after its end the wind begins to blow. The convective transport of the radioactivity clew arises. We trace the drifting over the terrain in the next four hours. Coincidental rain in the fourth hour of the convective transport can cause potentially dangerous increase of the deposited activity of 137Cs (red "patch" in Figure 3- Right).

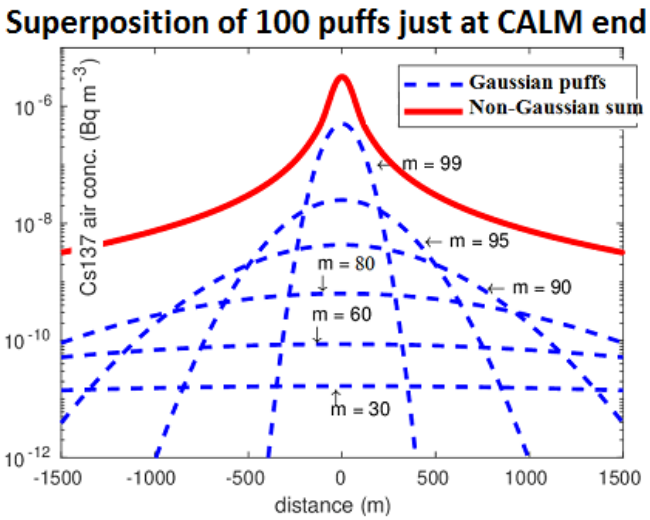


Figure 1: Blue lines: Gaussian distribution of ^{137}Cs concentration in air for a single puff m, each having its own age. Red line: Non-Gaussian distribution of the superposition of all puffs m; $m \in \{1,\ldots,M\}$.

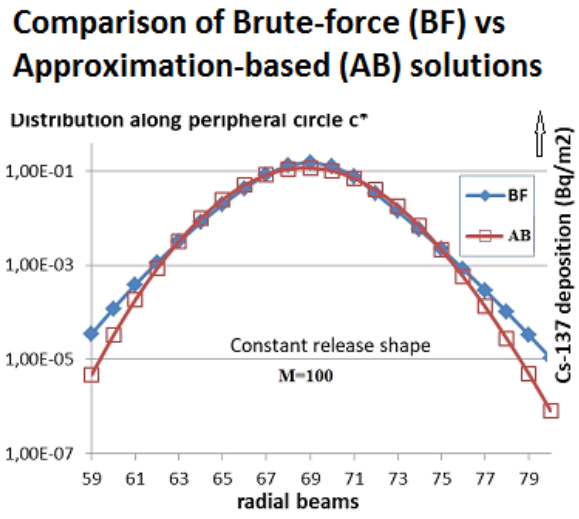


Figure 2: Comparison of two alternative transport BF and AB (in deposition of 137Cs on the ground around the circle c - see Figure 3, right).*

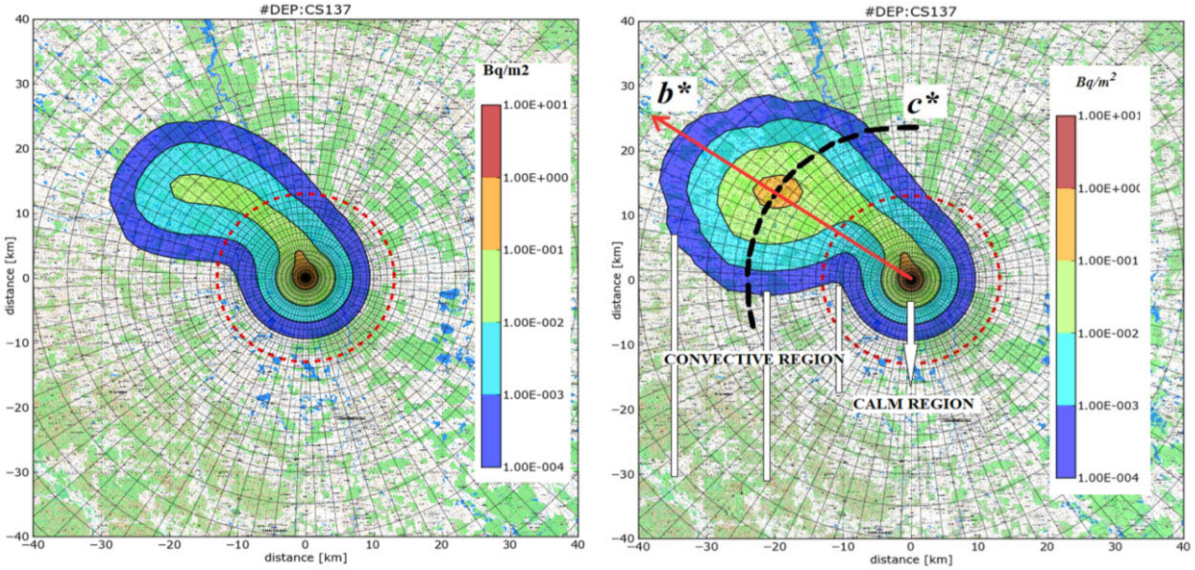


Figure 3: Deposition of radionuclide ^{137}Cs on terrain. M=100, BF solution, the situation just after 9 hours (5 hours of the calm plus 4 hours of the convective transport) from the beginning of the calm. Left: Without atmospheric precipitation. Right: The atmospheric precipitation 1.0 mm.h-1 in the fourth hour of the convective transport causes the serious increase of the deposited activity of ^{137}Cs.

Conclusion

Designed software offers a powerful tool for probabilistic analysis of consequences of the worst-case CALM scenario. The main contribution to methodology lies in proposing the novel approximation-based solution AB, its implementation and tests of its sufficient accuracy. It favourably affects three important areas:

a) AB approach provides comprehensible initial condition for the ensuing convective propagation in the form of optimal Gaussian approximant. The whole "super-puff" clue of radioactivity runs through the convective path only once. It opens way for potential applications of the authentic advanced dispersion codes.

b) Feasibility of prospective computationally intensive Bayesian methods of data assimilation. It has allowed the analysis exploiting the sampling-based procedure requiring a very high number of random samples.

c) Capability to perform the uncertainty and sensitivity analyses [3], forming a suitable base for advanced probabilistic consequence assessment. This probabilistic approach provides the essential data for the desirable transition from a deterministic procedure of the consequence assessment to the probabilistic approach, which enables generation of more informative probabilistic answers to assessment questions.

Links:

[L1] https://www.utia.cas.cz/

[L2] https://havarrp.utia.cas.cz/harp

References:

[1] P. Pecha, et al.: "Determination of Radiological Background Fields Designated for Inverse Modelling", J. of Atmospheric Environment, 246, 2021, 118105, 2020. http://library.utia.cas.cz/separaty/2021/AS/pecha-0537537.pdf

[2] M. Kárný M and P. Pecha: "Novel Simulation Technique of Harmful Aerosol Substances Propagation into the Motionless Atmosphere", Annals of Nuclear Energy, 165, 2021. https://doi.org/10.1016/j.anucene.2021.108686.

[3] P. Pecha and M. Kárný: "A new methodology is outlined and demonstrated on the improvement of uncertainty and sensitivity", Stochastic Environmental Research and Risk Assessment, 2021. https://doi.org/10.1007/s00477-021-02110-0

Please contact:

Petr Pecha, Institute of Information Theory and Automation, Czech Republic
pecha@utia.cas.cz

Formal Modelling and Optimal Traffic Management for Future Railways

by Francesco Flammini (Mälardalen University), Stefano Marrone (University of Campania "Luigi Vanvitelli") and Lei Chen (University of Birmingham)

PERFORMINGRAIL aims to delineate, through formal modelling and optimal traffic management, moving block railway signalling using advanced train positioning approaches for diverse market segments.

In order to address several challenges coming from the need to optimise their operational capacity, especially in highly congested corridors, future railways need to implement the new paradigms enabled by recent technological development, such as moving block signalling [1], satellite-based train positioning, and virtual coupling [2]. Those technologies have the potential to significantly increase the performance/cost ratio; however, they also pose several safety concerns that need to be formalised and carefully addressed by novel modelling and simulation approaches [3].

To tackle those challenges, the PERFORMINGRAIL project has been granted by the European Union Horizon 2020 framework program within Shift2Rail (S2R) Joint Undertaking (JU), Innovation Programme 2 (IP2) "Advanced traffic management and control systems", with a specific call addressing "Modelling of the Moving Block system specification and future architecture (TD2.3) + RAIM algorithms, Assessment Report and support for Railway Minimum Operational Performance Standards (TD2.4)" (call code: S2R-OC-IP2-01-2020).

The aim of the project is to implement a holistic system approach to address the open challenges for the Moving Block and Virtual Coupling concepts in terms of safe operational principles and specifications, high-accuracy train localisation and optimised moving block traffic management algorithms. The main objectives of the project are to enhance and verify existing specifications for moving block signalling, while developing formal models, algorithms, and proof of concepts to test and validate an integrated future moving block system architecture that will provide safe and efficient operational performances.

PERFORMINGRAIL will support the activities in the Shift2Rail IP2 Adaptable Communications Technology Demonstrator by helping to bring the developments as close as possible to the market while also helping to update the regulatory framework.

In accordance with the objectives of the IP2 TD2.4 Fail Safe Train Positioning (including GNSS) described in the S2R Multi Annual Action Plan (MAAP), in the field of Train Localisation (based on the use of combined technologies such as EGNSS, IMU, kinematics, Digital Map), PERFORMINGRAIL will:

- contribute Enhanced railways Fault Detection and Exclusion algorithms for addressing local feared events, and Data Fusion algorithms suitable for railway safe applications;
- contribute EGNSS monitoring techniques based on carrier phase measurement and multi-frequency technology
- help promote the use of formal methods to check safety of advanced railway operation (in accordance with IP2 TD2.7);
- provide Independent Assessment Report on the proposed technologies and solutions.

The project is structured in seven Work Packages, including five technical Work Packages (WP1 – WP5), plus WP6 about dissemination and exploitation, and WP7 for project administration and management.

Key technical objectives of the PERFORMINGRAIL project:

- OBJECTIVE 1: Definition of Specifications for safety and performance of moving block operations;
- OBJECTIVE 2: Formal modelling of moving block specifications;
- OBJECTIVE 3: Identification of moving block hazards;
- OBJECTIVE 4: Development of fail-safe train localisation solution;
- OBJECTIVE 5: Design of future traffic management architecture for moving block;
- OBJECTIVE 6: Implementation of a moving block testing platform;
- OBJECTIVE 7: Proof-of-concept and assessment of moving block specifications and models;
- OBJECTIVE 8: Recommendations for safety and performance of moving block configuratio;

In order to achieve those objectives, the research activities within PERFORMINGRAIL with be based on the results in cooperation with other S2R complementary projects such as X2RAIL-5 and X2RAIL-3.

Project goals can be achieved by adopting a model-driven and compositional modelling approach, which also naturally fits the EULYNX modelling method, as the formal models can be automatically derived from SysML and other high-level engineering specification languages through proper model transformations. The construction of the system models will be compositional, according to a bottom-up development methodology, enabling model parameterisation and consequently the possibility to easily adapt the models to different system configurations. So, model-driven techniques support automation, whereas compositionality allows for a major flexibility in modelling, promotes the reuse of models and facilitates the definition of modelling guidelines.

The reference architecture can be instantiated to define a proper tool chain also using standard languages and existing tools already employed by industries. It is characterised by three tiers, three actors and three layers (see Figure 1). The three actors represent different classes of users that may develop: domain-specific modeling languages (Language Engineer), transformations (Software Engineer) or models

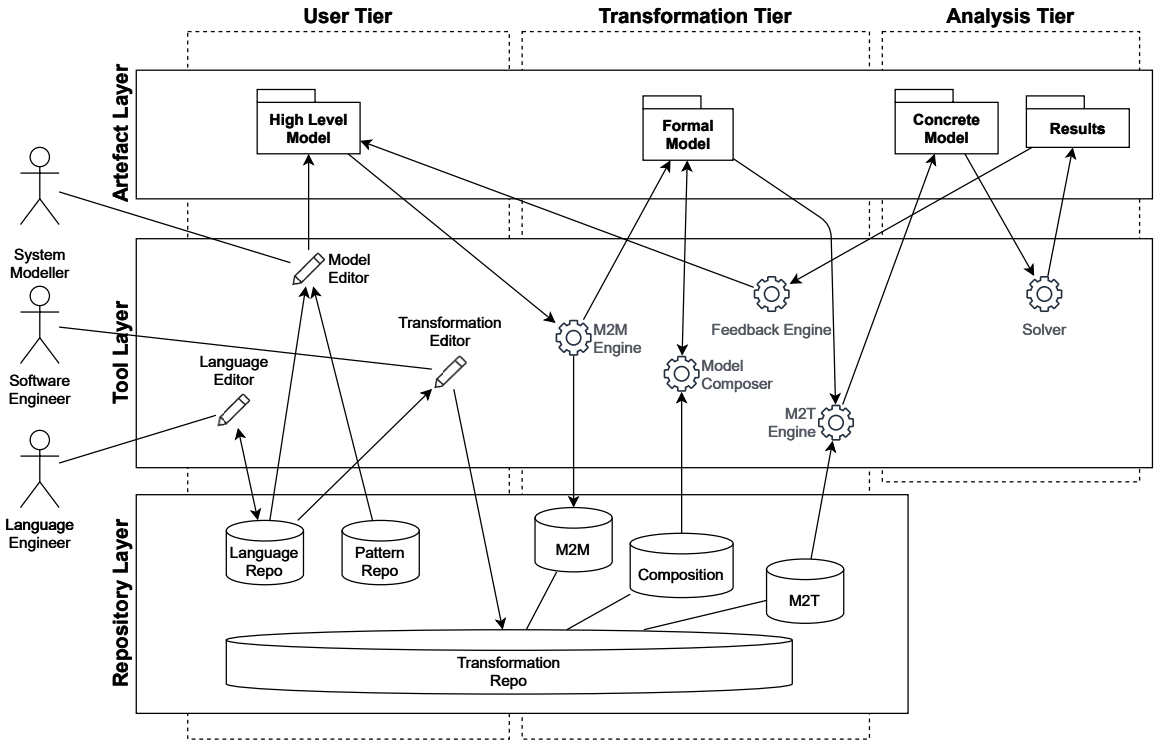


Figure 1: Reference architecture of the PERFORMINGRAIL modelling and analysis framework.

(System Modeler). The three horizontal layers (Artefact, Tool and Repository) represent the entities (models, tools and databases) involved in each tier. The three vertical tiers contain:

- Artefacts, tools and storage elements associated with non-automatic tasks (User);
- High-Level Models (e.g., UML models), intermediate formal models, Model-to-Model (M2M) and Model-to-Text (M2T) transformations and repositories associated with automatic tasks (Transformation);
- Final (concrete) models and tools for the analysis and the solution (Analysis).

In addition, a Model Composer is considered to compose sub-models into a single model if needed. The Feedback Engine is in charge of updating the High-Level Models after the analysis or test case generation phases.

PERFORMINGRAIL is coordinated by the University of Birmingham (UK), through its Birmingham Centre for Railway Research and Education (BCRRE), and project partners include TU Delft (NL), University Gustave Eiffel (FR), CINI (IT), Mälardalen University (SE), CERTIFER (FR), Rokubun (ES), and Eulynx (NL).

At the time of writing, the project has already achieved its first milestone (MS1) on schedule with the submission of deliverables:

- D1.1 Baseline system specification and definition for Moving Block Systems;
- D2.1 Modelling guidelines and Moving Block Use Cases characterisation.

Additional and up-to-date information about the project is available on the project website [L1] and social channels [L2], while technical project deliverables and research papers are also shared on ResearchGate [L3].

This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 101015416. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union.

Links:

[L1] https://www.performingrail.com/
[L2] https://twitter.com/PerformingRail
[L3] https://kwz.me/hfV

References:

[1] L. Carnevali, et al.: "Non-Markovian Performability Evaluation of ERTMS/ETCS Level 3", in Beltrán M., Knottenbelt W., Bradley J. (eds): "Computer Performance Engineering EPEW 2015", Lecture Notes in Computer Science, vol 9272, Springer, Cham.(2015) https://doi.org/10.1007/978-3-319-23267-6_4

[2] F. Flammini, et al.: "Compositional modeling of railway Virtual Coupling with Stochastic Activity Networks", Formal Aspects of Computing (2021) https://doi.org/10.1007/s00165-021-00560-5

[3] D. Basile, et al.: "Modelling and Analysing ERTMS L3 Moving Block Railway Signalling with Simulink and Uppaal SMC", in Larsen K., Willemse T. (eds): "Formal Methods for Industrial Critical Systems", Lecture Notes in Computer Science, vol 11687, Springer, Cham. (2019) https://doi.org/10.1007/978-3-030-27008-7_1

Please contact:

Lei Chen (Project Coordinator), The University of Birmingham, United Kingdom, l.Chen.3@bham.ac.uk
Francesco Flammini (Dissemination Manager), Mälardalen University, Sweden, francesco.flammini@mdh.se

SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Call for Proposals

Dagstuhl Seminars and Perspectives Workshops

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is accepting proposals for scientific seminars/workshops in all areas of computer science, in particular also in connection with other fields.

If accepted the event will be hosted in the seclusion of Dagstuhl's well known, own, dedicated facilities in Wadern on the western fringe of Germany. Moreover, the Dagstuhl office will assume most of the organisational/ administrative work, and the Dagstuhl scientific staff will support the organizers in preparing, running, and documenting the event. Thanks to subsidies the costs are very low for participants.

Dagstuhl events are typically proposed by a group of three to four outstanding researchers of different affiliations. This organizer team should represent a range of research communities and reflect Dagstuhl's international orientation. More information, in particular, details about event form and setup as well as the proposal form and the proposing process can be found on

https://www.dagstuhl.de/dsproposal

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is funded by the German federal and state government. It pursues a mission of furthering world class research in computer science by facilitating communication and interaction between researchers.

Important Dates

- *Next submission period:*
  April 1 to April 15, 2022
  Seminar dates: In 2023/2024.

Call for Papers

FM 2023: 25th International Symposium on Formal Methods

Lübeck, Germany, 6-10 March 2023

FM 2023 is the 25th international symposium in a series organized by Formal Methods Europe (FME), an independent association whose aim is to stimulate the use of, and research on, formal methods for software development. FME has a Memorandum of Understanding with the ERCIM Working Group on Formal Methods for Industrial Critical Systems (FMICS) to collaborate in holding an annual joint industry-focussed event. This Industry Day (I-Day) at FM targets the industrial development and use of formal methods.

Topics

FM 2023 will highlight the development and application of formal methods in a wide range of domains including trustworthy AI, software, computer-based systems, systems-of-systems, cyber-physical systems, security, human-computer interaction, manufacturing, sustainability, energy, transport, smart cities, healthcare, and biology. It particularly welcomes papers on techniques, tools and experiences in interdisciplinary settings. It also welcomes papers on experiences of applying formal methods in industrial settings, and on the design and validation of formal method tools.

New this year! FM 2023 explicitly welcomes submissions to the special FM 2023 session on "Formal methods meets AI", which is focused on formal and rigorous modelling and analysis techniques to ensuring safety, robustness etc. (trustworthiness) of AI-based systems.

Deadlines

The paper submission deadline is 11 September 2022, with an abstract due one week earlier. Accepted papers will be included in the Symposium Proceedings published in Springer's Lecture Notes in Computer Science series. Extended versions of selected papers will be invited for publication in a special issue of a journal.

General Chair

Martin Leucker (University of Lübeck, Germany)

PC Chairs:

- Marsha Chechik (University of Toronto, Canada)
- Joost-Pieter Katoen (RWTH Aachen University, Germany & University of Twente, The Netherlands)

More information: https://fm2023.isp.uni-luebeck.de/

ERCIM
European Research Consortium for Informatics and Mathematics

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Host of the World Wide Web Consortium.

Consiglio Nazionale delle Ricerche
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1, 56124 Pisa, Italy
www.iit.cnr.it

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
www.cwi.nl

RISE SICS
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
www.fnr.lu

SBA Research gGmbH
Floragasse 7, 1040 Wien, Austria
www.sba-research.org/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
www.ics.forth.gr

SIMULA
PO Box 134
1325 Lysaker, Norway
www.simula.no

Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
www.iuk.fraunhofer.de

Eötvös Loránd Research Network
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
www.sztaki.hu/

INESC
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal
www.inesc.pt

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
www.cs.ucy.ac.cy/

Institut National de Recherche en Informatique et en Automatique
B.P. 105, F-78153 Le Chesnay, France
www.inria.fr

Universty of Warsaw
Faculty of Mathematics, Informatics and Mechanics
Banacha 2, 02-097 Warsaw, Poland
www.mimuw.edu.pl/

I.S.I. – Industrial Systems Institute
Patras Science Park building
Platani, Patras, Greece, GR-26504
www.isi.gr

VTT Technical Research Centre of Finland Ltd
PO Box 1000
FIN-02044 VTT, Finland
www.vttresearch.com