Horizon 2020 Program (2014-2020)

Cybersecurity, Trustworthy ICT Research & Innovation Actions
Security-by-design for end-to-end security
H2020-SU-ICT-03-2018

Cyber security cOmpeteNCe fOr Research anD InnovAtion[†]

Work Package 1: European Secure, Resilient and Trusted Ecosystem (ESRTE)

Deliverable D1.2: 2nd Year Report on Designing and Developing an European Secure, Resilient and Trusted Ecosystem (ESRTE)

Abstract: This deliverable describes the research activities undertaken by work pack-age WP1 of the CONCORDIA Horizon 2020 project during the second year.

Contractual Date of Delivery	*M24*
Actual Date of Delivery	*30.12.2020*
Deliverable Dissemination Level	*Public*
Editors	*Mattijs Jonker (UT)* *Jean-Yves Marion (UL) Aiko Pras (UT)* *Jürgen Schönwälder (JUB)* *Nikos Salamanos (CUT)* *Michael Sirivianos (CUT)* *Marinos Tsantekidis (TUBS)* *Ramin Yazdani (UT)*
Contributors	*All partners involved in WP1*
Quality Assurance	*Thibault Cholez (UL)* *Detlef Houdeau (IFAG)*

[†]The research leading to these results has received funding from the European Union Horizon 2020 Program (2014-2020) under grant agreement n° 830927.

The CONCORDIA Consortium

CODE	Research Institute CODE (Coordinator)	Germany
FORTH	Foundation for Research and Technology - Hellas	Greece
UT	University of Twente	Netherlands
SnT	University of Luxembourg	Luxembourg
UL	University of Lorraine	France
UM	University of Maribor	Slovenia
UZH	University of Zurich	Switzerland
JUB	Jacobs University Bremen	Germany
UI	University of Insubria	Italy
CUT	Cyprus University of Technology	Cyprus
UP	University of Patras	Greece
TUBS	Technical University of Braunschweig	Germany
MUNI	Masaryk University	Czech Republic
BGU	Ben-Gurion University	Israel
OsloMET	Oslo Metropolitan University	Norway
ICL	Imperial College London	UK
UMIL	University of Milan	Italy
BADW-LRZ	Leibniz Supercomputing Centre	Germany
EIT DIGITAL	EIT DIGITAL	Belgium
TELENOR	Telenor	Norway
ACS	Airbus Cybersecurity	Germany
SECT	secunet Security Networks	Germany
IFAG	Infineon	Germany
SIDN	SIDN	Netherlands
SNET	SurfNet	Netherlands
CYD	Cyber Detect	France
TID	Telefonica I+D	Spain
RD	RUAG Defence	Switzerland
BD	Bitdefender	Romania
ATOS	Atos Spain S.A.	Spain
SAG	Siemens	Germany
Flowmon	Flowmon Networks	Czech Republic
T¨UV TRUST IT	TUV TRUST IT GmbH	Germany
TI	Telecom Italia	Italy
EFA	EFACEC	Portugal
ALBV	Arthur's Legal B.V.	Netherlands
EI	eesy innovation	Germany
DFN-CERT	DFN-CERT	Germany
CAIXA	CaixaBank	Spain
GSDP	Ministry of Digital Policy, Telecommunications and Media	Greece
RISE	RISE Research Institutes of Sweden AB	Sweden
Ericsson	Ericsson AB	Sweden
SBA	SBA Research gemeinnutzige GmbH	Austria
IJS	Institut Jozef Stefan	Slovenia
UiO	University of Oslo	Norway
ULANC	University of Lancaster	UK
ISI	Athena ISI	Greece
UNI PASSAU	University Passau	Germany
RUB	Ruhr University Bochum	Germany
CRF	CRF	Italy
ELTE	Eötvös Loránd University	Germany
Ultimaco	Ultimaco Management GmbH	Germany
RUAG	RUAG AG	Switzerland
EFACEC	Efacec Electric Mobility, S.A.	Portugal

Document Revisions & Quality Assurance

Revisions

Ver.	Date	By	Overview
0.1	2020-11-02	*M. Jonker (UT)*	Initial template
0.2	2020-12-22	*M. Jonker (UT)*	Internal review version
1.0	2020-12-29	*M. Jonker (UT)*	Final deliverable document

Contents

Executive Summary

The goal of CONCORDIA is to build a European secure, resilient and trusted ecosystem for innovation in the area of cybersecurity. It brings together partners from academia and industry to stimulate collaboration and increase the impact of European research. As stated in the Description of Action, the objectives for CONCORDIA's research Work Package (WP1) are: 1) to perform excellent academic research, 2) to organise scientific events, 3) to play a leading role in the organisation of the scientific community and 4) to contribute to standardisation, open data and code.

In last year's deliverable we could write that CONCORDIA researchers have produced more than 50 workshop, conference and journal papers. In this second year we are proud that more than hundred additional papers were written; see Appendix A for details. In fact, the total number of papers entered into the EU-ECAS system is even higher, since white papers and papers without peer-review are not counted in this deliverable. Although the objective regarding the number of papers, as stated in the Description of Action, has already been reached, also in the remaining years CONCORDIA researchers will continue to publish top papers. But not only the quantity of papers is high, but also quality, since several papers have been published at top venues, including USENIX Security, ACM-IMC, NDSS and IEEE INFOCOM.

In 2020 CONCORDIA researchers contributed to the organisation of 15 scientific conferences (see Appendix B) and acted as Technical Program Committee member for more than 70 different conferences (see Appendix C); like last year within CONCORDIA the most active persons in the organisation of events are women. CONCORDIA also organised together with the three other EU pilot projects (ECHO, SPARTA and Cybersec4Europe), the 4th International NeCS PhD Winter School (Trento) and the IFIP Summer School on Privacy and Identity Management (virtual conference). A novel initiative in 2020 was the organisation of the 'Early Stage PhD workshop' and the CODE 'research day', also aimed at PhD education.

Again this year, CONCORDIA researchers played a leading role in the organisation of the scientific community by acting as chairs of the IFIP Technical Committee 6 (Communication Systems), IFIP Working Group 6.6 (Management of Network and Distributed Systems), and as editorial board members for 25 journals and transactions, ranging from from ACM, IEEE, Springer and Wiley.

Next to the traditional academic venues, in 2020 CONCORDIA researchers also contributed to events for governments, industry and the wider Internet community (network operators, Internet engineers and architecture designers). A concrete example is task T1.2's investigation into DNS zone administration errors, which resulted in presentations at DNS-OARC and RIPE80, among others, and prompted changes by DNS operators to improve DNS security and stability. Other exam-

ples are the white papers by T1.5 that study the impact of Twitter and YouTube on the US elections, and the blog on 'Ethics, Responsibilities and Vulnerabilities' at FIRST.

During the first year of the project, collaboration between WP1 tasks and the pilots was explored. Possible synergetic research activities were chartered. Over the second year, some of the planned work came to fruition. A concrete example is the joint work between network-centric security research (T1.2) and the DDoS Clearing House for Europe pilot (T3.2). Actionable intelligence from T1.2 can now feed into the pilot's architecture to enrich DDoS attack signatures. Another example is the strong collaboration that has been established between application-centric security research (T1.4) and the Threat Intelligence for the Telco Sector pilot (T2.1), where T1.4's efforts in investigating application security has been instrumental in making sure that requirements set by T2.1 are met.

As part of the task specific results, this deliverable includes security roadmaps for devices, networks, systems, applications and users. The starting point is digital sovereignty, which is the societal challenge that Europe is currently facing. One main outcome of these roadmaps is that additional research is needed to design a 'Responsible Internet', a novel security-by-design extension of the Internet that offers users better grip on dependencies and that increases trust. The notion of a Responsible Internet goes much further than 5G security, which focusses primarily on the cellular access part of the Internet. Note that the complete CONCORDIA roadmap can be found in Deliverable 4.4.

The second year of the project coincided with the outbreak of the COVID-19 pandemic. The positive news is that the immediate impact on WP1 activities has been relatively limited, although not nil. The negative news is that the long-term effect on the mental wellbeing of particularly young researchers, who often live abroad where they still have to built a social network, is unknown. Many of them live quite isolated and we may expect that for some the effects will be serious. The pandemic has also made attending events in-person virtually impossible, which hinders the generation of new ideas and initiativies.

In general we may conclude that WP1 has at least reached, and with respect to the number of publications, again exceeded the objectives as described in the Description of Action. Collaboration between research in WP1 and the pilots in WP2 and WP3 was strengthened, and more attention was paid to create impact by presenting at non-academic events. Although this worked quite well for some research activities, for others there is still potential for improvement, although it should also be said that due to COVID a number of activities could not take place.

1 Introduction

This document reports the research activities undertaken by the work package WP1 of the CONCORDIA project during the second year of the project. The goal of WP1 is to organize and coordinate scientific research within the CONCORDIA project. WP1 has the following objectives:

- Excellent academic research to build an European Secured, Resilient and Trusted Networked Ecosystem, papers for scientific journals, conferences and workshops;
- Organization of scientific events in the area of cybersecurity, including a dedicated annual European cybersecurity conference;
- Leading role in the organization of the scientific community, outreach to different target audiences, including public media and the general public;
- Contributions to standardization, open research data and code, shared via systems such as GitHub.

The main objective is to stimulate the publication of scientific results in key journals, conferences, and workshops in the broad field of cybersecurity. The SMART objective is to publish at least 100 of such papers during the project's lifetime.

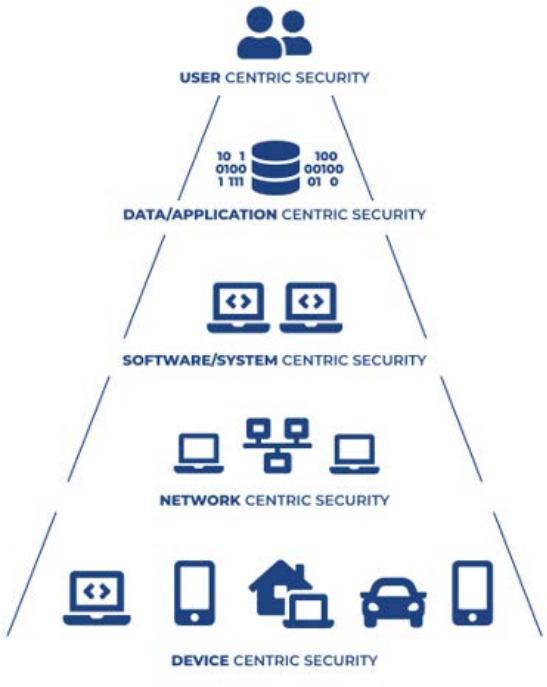

Figure 1: Tasks of WP1

The research under WP1 is organized into five tasks, with each one focusing on a particular aspect of cybersecurity (see Figure 1):

- T1.1: Device security aspects
- T1.2: Network security aspects
- T1.3: Software and system security aspects
- T1.4: Data and application security aspects
- T1.5: User security and privacy aspects

1.1 Structure of the document

This deliverable is structured as follows. Section 2 outlines key achievements of the second year. The next section, Section 3, discusses what actions have been taken after both reviews earlier in 2020. For each recommendation of the review report it provides an answer.

Thereafter, the research activities of the various tasks are discussed in detail. An overview of research related to device-centric security is provided in Section 4. Section 5 outlines network-centric security research activities. Section 6 discusses research concerning software and system-centric security aspects. Data and application specific security aspects are the focus of Section 7. Finally, Section 8 covers efforts related to user security and privacy.

For each of the five research tasks a similar structure is followed. Each task starts with shortly summarising its goals, followed by an introductory text that identifies the highlights and a flash forward. After that, the relation and collaboration with the pilots in WP2 and WP3 are detailed.

As a reaction on previous requests by the reviewers and the EU, we have included for each task ”considerations” for a security roadmap. Note that the text in these sections is largely equivalent to the text in Deliverable 4.4 (Cybersecurity Roadmap for Europe). In addition, each tasks has also included a subsection to detail which consequences of COVID-19 were specific for this task.

After these common roadmap and COVID-19 subsections, each task continues describes its specific research details. This deliverable concludes with a summary of the organization of scientific events and the scientific community in Section 9. Section 10 discusses the contributions to standards and open research data. The main text of the deliverable concludes with Section 11 that provides the conclusions and an outlook for 2021. The appendixes give details regarding publications, conferences, technical program committee (TPC) memberships as well as editorial boards.

1.2 Impact of COVID-19 on Research Activities

2020 has been a special year. It seems that most of the research by our PhD students was not directly affected by COVID-19, and the writing of scientific papers could continue according to plan. In fact, it could even be that a short-term effect of COVID-19 is that PhD students could spend more time on writing papers, and that we'll see temporarily a higher output of papers.

Within WP1 most researchers do not need physical access to large laboratory infrastructures, except for some cases within T1.4 and T1.5 In T1.4, the Machine Learning Platform located in Oslo suffered from a severe malfunction, which hindered further research activities using that platform. In T1.5, the research to be performed by Telefonica I+D (TID) was effected, since it was not possible to perform the regular research internships and get access to TID's laboratory and other resources. For details, see the task specific sections of this deliverable.

As a consequence of not being able to travel, most tasks within WP1 moved to bi-weekly or monthly online meetings. Such meetings seem to work well.

However, the real impact of COVID-19 will only become apparent later. Many young PhDs pursue their PhD in another country, far from their family and friends. They often live in small student houses, with limited possibilities 'to escape". The long term impact on their well-being is hard to predict, but we should expect that for some the consequences will be severe.

Due to COVID-19 the possibilities to meet new people and generate new ideas and take new initiatives have become very limited. The long-term effects of this can not yet precisely be estimated, but will likely be serious.

Within this deliverable we have included dedicated sections per task to describe the impact of COVID-19 on their specific research activities.

2 Key Achievements

Excellent Academic Research

The main objective of WP1 is to stimulate the publication of scientific results in the broad field of cybersecurity (see Section 1). As we reported previously in deliverable D1.1, about 50 papers had been published in workshops, conferences and journals during year one of the CONCORDIA project. This was a good indicator that the SMART objective to publish at least 100 of such papers during the project's four-year lifetime would be met. We also pointed out in D1.1 that due to the project's young age at the time, combined with the fact that publishing top research takes time, we expected the number of papers published at top venues to increase over the years to come. This developed as expected. During the second year, a little over 100 additional papers were published (see Appendix A). Although this SMART objective has therefore already been met, CONCORDIA partners will continue writing high quality papers in the next two years.

Organization of Scientific Events

During the second year of the project, CONCORDIA researchers were member of the technical program committees (TPC) of over 70 different conferences (nearly double that of year one). CONCORDIA researchers also contributed to the organization of nearly 15 scientific conferences (down from about 25 during year one). Additional details are provided in Section 9.

In year one, CONCORDIA started efforts to co-organize, together with the three other EU pilot projects (ECHO, SPARTA and Cybersec4Europe), two PhD schools. These ongoing efforts were previously reported on in deliverable D1.1. The PhD schools, specifically, are the 4th International NeCS PhD Winter School, and the 15th IFIP Summer School on Privacy and Identity Management. The PhD Winter School took place in Trento, Italy, from January 20 to 24, 2020. It was successful and well-received by participants. Due to the COVID-19 outbreak, the Summer School on Privacy and Identity Management had to be rescheduled to September 2020 as a fully virtual event. It was a success nonetheless.

Organization of the Scientific Community

Professional organizations play an important role in organizing the scientific community. As in year one of the project, various CONCORDIA researchers held positions in such organizations during the second year. A CONCORDIA researcher chairs the IFIP Technical Committee 6 (TC6). TC6 focuses on *Communication Systems* and is the largest TC within IFIP. The chair of Working Group 6.6 of TC6 is also a CONCORDIA member. WG 6.6 focuses on the management of network and distributed systems.

A number of CONCORDIA researchers have sat on journal editorial boards during year two and are members of steering committees. A total of 25 journals by publishers such as ACM, Springer and IEEE are involved. Further details are (too) given in Section 9.

Standardization Efforts and Contributions to Open Tools and Research Data

The fourth objective of WP1 (see Section 1) is to contribute to standardization efforts, open research data, and tools/software.

During the second year of the project, CONCORDIA researchers worked on two Internet RFCs that pertain to securing networked applications. Four open data sets were produced, account for diverse, security-related traces such as traffic from IoT devices, TLS v1.3 deployment measurements, and malicious cryptocurrency mining code.

Various open-source tools were released by CONCORDIA researchers during the second year. A notable example is *LoRadar*, which is a tool that enables investigations into the performance, utilization, and security aspects of Long Range (LoRa) networks (LoRa is a Low Power WAN technology).

More details can be found in Section 10. Note that CONCORDIA knows two tasks (T5.3 and T6.4) that relate specifically to standardization and open data. As a consequence, various CONCORDIA efforts related to standardization and software are not reported on in this deliverable. An extended overview of activities can be found in the deliverables of WP5 and WP6.

3 Recommendations from the review

The Section discusses the recommendations provided by the reviewers in their first (January 2020) and second (September 2020) review reports, as far as these comments relate to WP1.

3.1 1st review

Feedback: *Collaborative approach and tools/formats chosen are appropriate and allow for amplification of the effort. However, to achieve the academic-industry knowledge transfer that CONCORDIA has set out for, seeking more applied and industry fora would be encouraged. This is particularly true for publications (where academic work can be reformatted for industry or general publication) as well as events (in particular seeking out less academic, government technical /industry conferences like CyCon or FIC).*

Answer: We have taken a number of actions, but we also realise that such actions need to be continued throughout the project's lifetime, especially since young researchers tend to focus their work on their own detailed problems and sometimes forget about the bigger picture.

The reviewer's feedback was discussed with the WP1 researchers, and we asked them to invest also in non-academic events. From their reactions it became clear that several of them already present, attend and organise non-academic events at a regular basis. Below we'll provide an overview.

Presentions: CONCORDIA researchers provided or contributed to presentations at the following 'non-academic' events.

- **ICANN**[1]
  Several researchers attended the ICANN (The Internet Corporation for Assigned Names and Numbers) 68 and 69 meetings. At ICANN 68 Cristian Hesselman [SIDN] organised a plenary session on the interplay between the IoT and the DNS and was one of the speakers[2]. At the DNSSEC Workshop of ICANN 69 Moritz Müller [SIDN/UT] presented the paper titled: *The Reality of Algorithm Agility: Studying the DNSSEC Algorithm Life-Cycle*; this paper was originally presented at the ACM Internet Measurement Conference [99].
- **IM'21**[3]
  A number of PhD students got already papers accepted at the "experience session" of IM'21, which goal is to *'complement the Technical Sessions with contributions that emphasize practical experiences and lessons learned in*

[1] https://www.icann.org
[2] https://www.thnic.or.th/en/icann68plenarysession/
[3] https://im2021.ieee-im.org/call-experience-session-papers

applying, implementing, and deploying management technologies. ... Experience Sessions are particularly aimed at decision makers and experts from industry' .

- **RIPE**[4]
  Raffaele Sommese [UT] and Giovane Moura [SIDN] attended the RIPE 80 event. Raffaele Sommese gave a talk titled: *When Parents and Children Disagree: Diving into DNS Delegation Inconsistency (RACI)*[5]. Giovane Moura was co-author of the paper: *Counterfighting Counterfeit: detecting and taking down fraudulent webshops at a ccTLD*, which was presented by his team member from SIDN[6].
- **GTER/GTS**
  Leandro Bertholdo [UT] gave a presentation at the Brazilian GTER 49 — GTS 35 (Internet Infrastructure Week in Brazil) meeting with the title: *BGP Anycast Tuner: Intuitive Route Management for Anycast Services*[7].
- **DNS-OARC**[8]
  Raffaele Sommese, Giovane Moura and Cristian Hesselman attended the DNS-OARC 33 event, which is organised by the DNS Operations, Analysis, and Research Center. Raffaele Sommese presented: *The Forgotten Side of DNS: Orphan and Abandoned Records*[9]. Giovane Moura and Cristian Hesselman were co-author of the paper: *Clouding up the Internet: how centralized is DNS traffic becoming?*, which was presented by a colleague from UCLA (Wes Hardaker) [10].
- **SURF**[11]
  In February 2020 Cristian Hesselman presented at the SURF Security and Privacy Conference a paper titled *Increasing the Netherlands' DDoS resilience together.*[12]
- **ETNO**[13]
  Also in February 2020 Jelte Jansen (SIDN) gave a presentation at the ETNO Working Group Meeting on the subject *The IoT and the DNS*".[14].

[4] https://www.ripe.net
[5] https://ripe80.ripe.net/archives/video/324/
[6] https://ripe80.ripe.net/archives/video/322/
[7] https://ftp.registro.br/pub/gter/gter49/04-BgpAnycastTuner.pdf
[8] https://www.dns-oarc.net
[9] https://indico.dns-oarc.net/event/34/contributions/794/
[10] https://indico.dns-oarc.net/event/34/contributions/790/
[11] https://www.surf.nl/agenda/surf-security-en-privacyconferentie
[12] https://www.sidnlabs.nl/downloads/23YVgl8Zv90T1bJFR28SXu/58a9485892e2fb4cdba40a1b6a341d57/20200207-SURF-ddos-clearing-house-final.pdf
[13] https://etno.eu
[14] https://www.sidnlabs.nl/downloads/3mLIZRToCec91xGGxDiRxv/a5deab780b5a0e4509f9aa530f2bc6f7/The_IoT_and_the_DNS.pdf

- **PQC**[15]
  At the PQC NIST Virtual Workshop on 'Considerations in Migrating to Post-Quantum Cryptographic Algorithms' in October 2020, Moritz Müller presented *Retrofitting Post-Quantum Cryptography in Internet Protocols: A Case Study of DNSSEC*[16].

As becomes apparent from the list above, not all CONCORDIA partners present (yet) at 'non-academic' events. One limiting factor is that, before researchers submit to an 'unknown' conference, they generally attend such conference once or twice to better understand 'the DNA' of that conference. 'Submitting in the wild' is what many researchers like to avoid. Unfortunately, due to COVID-19, travelling to novel conferences and meeting with people in person and "on local time" has been impossible, which limited the attendance at new conferences.

Attendance: Several CONCORDIA researchers, and in particular the more senior researchers, attend different kinds of 'non-academic' events. Since such attendance is often not reported in academic research reports, below is therefore just a small subset of such attendance.

- **CES**[17]
  Researchers from CYD joined the January 2020 CES conference in Las Vegas.
- **FIC**[18]
  Researchers from UL joined the January 2020 Forum International de la Cybersécurité (FIC) in Lille (France).
- **Quarkslab**[19]
  In October 2020 researchers from UL and CYD joined the webinar of the company Quarkslab about the Gorille Cloud (see Section 6 for further details).
- **FIRST**[20]
  Jeroen van der Ham (UT) joined the November 2020 Virtual First conference.
- **ONE**[21]
  Aiko Pras (UT) joined the November 2020 ONE conference, organised by the Dutch national Cyber Security centre.

[15] https://https://www.nccoe.nist.gov/events/virtual-workshop-considerations-migrating-post-quantum-cryptographic-algorithms
[16] https://www.sidnlabs.nl/downloads/6VtmFHbZubErC0Zibh83gZ/5a5e40df242e670ff5c70b336ab08c14/Mueller-DNSSEC-PQC_workshop_long.pdf
[17] https://www.ces.tech/
[18] https://2020.forum-fic.com/accueil.htm/
[19] https://quarkslab.com/
[20] https://www.first.org/conference/2020/
[21] https://one-conference.nl

- **360-Grand-Est**[22]
  In December 2020 Jean-Yves Marion (UL) attended and presented at the cyber security event organised by Région Grand-Est (France).[23]
- **ANRW**[24]
  A number of CONCORDIA researchers will join the ACM/IRTF Applied Networking Research Workshop (ANRW) 2021, which *'provides a forum for researchers, vendors, network operators, and the Internet standards community to present and discuss emerging results in applied networking research'*. Roland van Rijwijk (UT) is Steering Committee member of that workshop.
- **DINR**[25]
  Mattijs Jonker (UT) joined the July 2020 DNS and Internet Naming Research Directions 2020, to discuss the research landscape around DNS and Internet naming. This workshop was organised by the University of Southern California Information Sciences Institute (USC/ISI).

White paper and blogs: Several CONCORDIA partners wrote white papers and blogs. Examples are:

- *Ethics, Responsibilities, Vulnerabilities* , a blog coauthored by Jeroen van der Ham (UT) and published by FIRST, the Forum of Incident Response and Security Teams, https://www.first.org/blog/20200518_Ethics_Responsibilities_Vulnerabilities
- *On the Influence of Twitter Trolls during the 2016 US Presidential Election*, by Nikos Salamanos, Michael J. Jensen, Xinlei He, Yang Chen and Michael Sirivianos (CUT), https://arxiv.org/pdf/1910.00531.pdf
- *Understanding the Incel Community on YouTube*, by Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburny, Emiliano De Cristofaroz, Gianluca Stringhini and Michael Sirivianos (CUT), https://arxiv.org/abs/2001.08293
- *Did State-sponsored Trolls Shape the US Presidential Election Discourse? Quantifying Influence on Twitter*, by Nikos Salamanos, Michael J. Jensen, Xinlei He, Yang Chen, Costas Iordanou and Michael Sirivianos (CUT), https://arxiv.org/abs/2006.09938
- *"It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations*, by Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini and Michael Sirivianos (CUT), https://arxiv.org/pdf/2010.11638.pdf
- *A responsible internet: increasing trust in the foundation of digital societies* , by Cristian Hesselman (SIDN), https://www.sidnlabs.nl/en/new

[22] https://www.360grandest.fr
[23] https://www.360grandest.fr/session/1bf3ff1a-e912-eb11-9fb4-0003ff1d36cc
[24] https://irtf.org/anrw/2021/
[25] https://ant.isi.edu/events/dinr2020/

s-and-blogs/a-responsible-internet-increasing-trust-in-the-foundation-of-digital-societies

- *MAnycast*^{2}*: Using anycast to measure anycast*, by Raffaele Sommese (UT), https://blog.apnic.net/2020/12/15/manycast2-using-anycast-to-measure-anycast/

Workshop organisation:

One effective measure to improve the academic-industry knowledge transfer, is to increase the number of PhDs that perform research within the CONCORDIA ecosystem. To get more PhDs involved, we have organised a number of workshops and activities that focussed on early stage PhDs.

One activity we organised on June 18 2020, was a workshop for Early Stage PhDs, where such starting PhDs presented their research plans. They obtained feedback not only from senior scientists, but also from the industrial partners within CONCORDIA. In conjunction with the CODE Jahrestag[26] we organised a Science Track, at which PhDs presented their work to Governmental and Industrial partners.

Feedback: *Whereas the different research activities concluded by the academic and research community are well detailed and an important work has been concluded for each of the five cybersecurity aspects, it does not emphasise sufficiently on the interactions between the different tasks. A number of meetings have been organised to allow participants to make acquaintance of each other. However, as the main goal of WP1 is to design and develop an ecosystem in Cybersecurity, the collaboration between the researchers involved in the different research domains is not sufficiently explained and shown.*

Answer: In the Task specific sections of this deliverable the various Task Leaders have detailed their interactions with the CONCORDIA Pilots of WP2 and WP3. Several workshops were organised between WP1 and the CONCORDIA Pilots. On May 27 as well as June 9 2020 workshops were organised between WP1 and T2.1. On June 16 an Industry symposium was organised by T1.5. On November 25 a workshop 'DDOS/Flooding attacks on mobile network Workshop' was organised between WP1, T2.1 and T3.2.

Feedback: *To have a harmonised structure between all the deliverable add a subsec entitled Structure of the document in the introduction.*

Answer: We included such subsection in the Introduction of Deliverable 1.2.

Feedback: *In tables of the Annex C and D, order them according to the name of the conferences/ journals and for each conference or journal, put the name of the researchers involved + the position name.*

[26] https://www.unibw.de/code-events/program-code2020

Answer: In this deliverable we have now ordered the entries of appendices B, C and D alphabetically. Each entry is followed by the name of the researcher(s) involved, their affiliation and role where applicable.

3.2 2nd review

Feedback: *The publication and event output is impressive and greatly reflects the ability and diversity of CONCORDIA participants. In the next review period, the project should further expand on if/how/whether the consortium has facilitated or encouraged joint research. Are there publications or events that would not have happened if there was no Concordia?*

Answer: WP1 has active collaboration with the pilots in WP2 and WP3. An example of such collaboration is on the topic of DDoS and 5G, for which multiple joint meetings as well as a workshop were organised together with T2.1 and T3.2. It must be noted, however, that COVID somehow prohibited the establishment of new collaborations, which often start at physical meetings after informal talks during coffee time. For example, a visit of young researchers to DFN in Hamburg had to be cancelled due to COVID-19.

Four partners within T1.3 (BD, CYD, ISI and UL) are addressing the challenge in the fight against malware. This research is the beginning of a collaboration between the two companies (BD and CYD), on one hand and two academic players (ISI and UL). The development of Gorille Cloud made jointly by CYD and UL is a promising solution to tackle cybercrimes. Within that context there is also collaboration on machine learning methods for malware analysis, which can be considered as emerging solutions, but which have yet to show their capacity in real situations.

Another highlight is the works of ULANC and RISE partners addressing the challenging of investigating vulnerabilities in complex IT environment including Cloud infrastructure and IOT devices.

Feedback: *Does Concordia have these results because it has picked participants well or has the consortium pushed the participants to go further?*

Answer: It is hard to prove that collaboration and paper publication would not have happened without CONCORDIA, but it seems very likely that several joint papers would not have been written without CONCORDIA. Examples of organisations that have written in recent months joint papers include ACS, FORTH, GSDP, SIDN, UI, UMil, CODE and the UT.

A good example that can be emphasised is the interaction between T1.2, T2.1 and T3.2 on DDoS protection. As demonstrated in Appendix A, this interaction has resulted into multiple joint papers between researchers from the UT (T1.2) and SIDN (T3.2): [67][99][98][17][96][145][66].

Feedback: *The approach of Concordia to empower young colleagues/ PhD students is systematic and focused. Has the consortium considered further how to make sure this stick thought their career trajectory, make the current work sustainable well after the consortium? There is such clear potential for these relationships to last through the professional lives of the current young researchers as they are building their professional networks.*

Answer: This is certainly on the radar of the CONCORDIA management, but is at this stage still a bit early to decide upon. One aspect that is unclear, is what the EU plans are regarding the four pilot projects. It is expected that joint events, such as the Winter and Summer Schools, will continue in some way, but for a decision on if or how to continue these joint events it seems to be still a bit early.

For early stage PhDs CONCORDIA has organised two events, one of them in conjunction with the CODE Jahrestag. It is expected that the event in conjunction with the CODE Jahrestag will remain to exist after the CONCORDIA project.

For PhD students an instructional video has been created on how to write a scientific paper [27]. That video is hosted on YouTube, and will remain to be available after the project ends.

[27] https://www.youtube.com/watch?v=5zthkvzyTfk

4 Device-Centric Security (T1.1)

Task 1.1 (T1.1) of the CONCORDIA project is concerned with device-centric security, with a special focus on IoT devices with limited resources. The task aims at

- developing techniques for detecting misbehaving IoT devices,
- analyzing automated software update mechanisms of IoT devices,
- investigating hardware components for post-quantum cryptography, and
- researching code analysis techniques to detect advanced persistent threats.

4.1 Overview

The research untertaken in 2020 can be classified into the four broad topics (i) attack prevention, (ii) utilizing trusted execution environments, (iii) device identification and security assurance, and device specific (iv) data and communication security.

- **Attack Prevention**

  Section 4.4.1 summarizes research on the efficient implementation of post-quantum cryptography on embedded devices and the applicability of machine learning techniques for more efficient side-channel analysis. Section 4.4.2 more generally addresses on the question how architectures for secure CyberPhysical systems can be found.

  The work described in Section 4.4.3 aims at providing a secure comprehensive authentication framework for users as well as for IoT devices. Finally, Section 4.4.4 outlines research on instruction set randomization that aims to prevent code injection attacks.

- **Trusted Execution Environments**

  Trusted execution environments such as ARM's TrustZone or Intel's Enclaves are the focus of a number of research efforts. Section 4.4.5 describes work assessing software utilizing TrustZone-M. Section 4.4.6 more specifically looks at secure communication between the normal and secure world while the research described in Section 4.4.7 aims at finding efficient ways of detecting code-reuse attacks on ARM Cortex-M processors.

  The work summarized in Section 4.4.8 uses Intel's Enclave technology to verify operating system integrity. In a similar way, Section 4.4.9 describes an approach enabling privacy-preserving malware analysis in the cloud. Finally, Section 4.4.10 describes an infrastructure enabling the parallel execution of JavaScript code on multiple remote nodes utilizing Intel SGX enclaves.

- **Security Assurance**

  The research summarized in Section 4.4.11 is exploring machine learning techniques for device identification purposes. These techniques can be used to identify devices based on external observations. A different approach to identify devices is using physically unclonable functions. Research in this direction is described in Section 4.4.12.

  Section 4.4.13 more broadly looks at security assurance technology for IoT systems. It stresses the importance of verifiable trustworthy data. Finally, Section 4.4.14 describes an implementation of the RESTCONF protocol for resource constrained embedded Linux systems. It can be used to prototype remote attestation mechanisms currently being standardized.

- **Data Security**

  A block-chain based architecture to store data in industrial IoT systems is presented in Section 4.4.15. A long range (LoRa) cloud radio access network using software-defined radios is described in Section 4.4.16.

  The research outlined in Section 4.4.17 aims at enforcing privacy preferences on IoT systems that use the MQTT protocol to communicate sensor readings to a data collection platform. A special aspect is the identification of emergency situations, during which less strict privacy preferences may be enforced. Finally, Section 4.4.18 discusses a technique for privacy preservation in industrial IoT systems where noise is used to prevent unauthorized access to data.

Figure 2 indicates how research in task T1.1 relates to the CONCORDIA pilots.

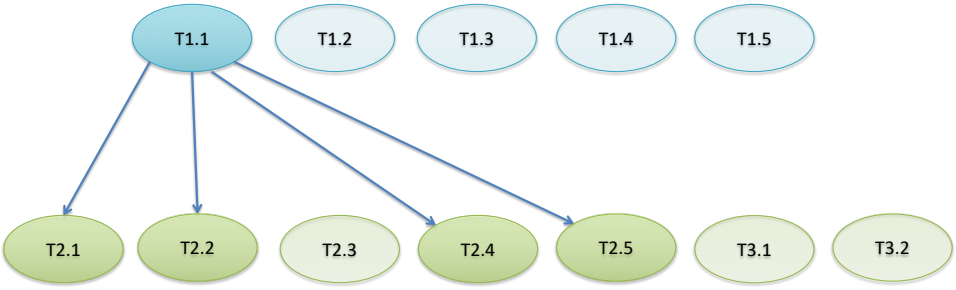

Figure 2: Links between task T1.1 and the CONCORDIA pilots

The research on the security assurance of IoT systems has direct relation to the Telecom Sector Pilot (Task T2.1) as explained in Section 4.4.13. The work on password-less authentication (Section 4.4.3) is related to the Finance Sector Pilot (Task 2.2). Finally, the security assessment of software based on TrustZone-M (Section 4.4.5) is relevant for the e-Health Sector Pilot (Task 2.4) and the Defense Sector Pilot (Task T2.5).

4.2 Security Roadmap Considerations

An activity has been started in Fall 2020 to identify research challenges that will be important for achieving digital sovereignty and an increased level of information technology security at the European level. The objective is to identify research challenges that can act as enablers for the European industry to build the most secure products in the world (Security made in Europe). Some research challenges will be more short-term, others may be medium- to long-term.

The need to improve the security of devices is to a large extend motivated from the dramatic growth of the Internet of Things (IoT). As part of their home automation, end-users will connect tens of billions of consumer devices to their Internet. To protect the privacy of these end users and to avoid that these devices become part of a botnet, security awareness and measures should be strengthened. Less visible, but from a digital sovereignty point of view probably more important, are the devices that are embedded within cars, drones and the devices that control our critical infrastructures and industrial systems.

To ensure Europe's digital sovereignty, it is essential that Europe keeps its ability to develop its own hard- and software infrastructures. In the past, Europe always had a strong chip industry, and for the future we should ensure that Europe remains the ability to design and manufacture its own high-performance microprocessors and other chips. In the next decades we may expect that traditional computers will partially be replaced by Quantum Computers, which implies that Europe should strengthen its research in the area of Quantum Computers.

4.2.1 Certified Software Update Mechanisms

Even if devices are tested and certified to be secure, vulnerabilities will be discovered sooner or later. It is therefore important that each device includes facilities to be updated. Software needs regular maintenance and regular security updates. Even if devices are tested and certified to be secure, vulnerabilities will be discovered sooner or later. It is therefore important that all device includes facilities allowing the software of the devices to be updated. This is even more important for embedded devices that are often not explicitly managed or that are operated by users that have little technical know-how or time available to care about the software in embedded devices.

Furthermore, many embedded devices have rather long lifetimes, during which companies producing products may cease to exist or they may be taken over by other companies. It is therefore necessary to develop secure software update mechanisms that not only protect and automate software updates but that in addition take into account the different possible product life-cycles. Solutions must be able to enforce updates and they must be able to disable updates (in case certain updates are known to be vulnerable). For some deployments, clearance of updates by third parties may be required, e.g., through a certification process of a European agency.

Such a certification process should foresee mechanisms of re-certification in order to deal with new attack techniques that are developed during the lifetime of deployed embedded devices.

Actions:

- Develop software update solutions that support long and complext device life-cycles during which responsibility for device updates may move between organizations.
- Develop, standardize and deploy secure software update mechanisms that support the certification of updates by some third party, e.g., the European Cybersecurity Agency ENISA.
- Develop an automated re-certification solution, whereby devices can be issued with an EU-backed security certification that is valid until a vulnerability is discovered. When this occurs, devices must be patched and re-certified without any physical interaction.

4.2.2 Device Identification and Assessment Mechanisms

Secure device identification is an essential step for establishing trust into a distributed computing environment. Being able to distinguish a clone from an expected genuine device is essential but not trivial. One approach is to design hardware components that can safely store device identity information (e.g., a device key) such that it is impossible to clone the stored information. The current trend is to make these hardware components more flexible and programmable, which leads to a situation where the complexity of the security software grows to a point where the correctness of the security software cannot be guaranteed anymore. An alternative approach is to use physically unclonable properties of a device to establish the identity of the device.

Related to the identification of the device is the identification of the software components that are installed and/or running on a device. It is necessary to continuously assess the integrity of the software components and to detect attempts to compromise a device, including attacks exploiting so call zero-day vulnerabilities.

Device identification and assessment mechanisms need to be complemented by remote attestation protocols, which enable authorized third parties to assess the integrity of a device and its software and to detect changes. These protocols should be standardized, and the industry will benefit from openly available reference implementations.

Actions:

- Develop device identification mechanisms that exploit physically unclonable properties of devices.

- Develop novel techniques to continuously assess the integrity of installed and running software and that can detect deviations from expected normal control flows.
- Create and extend standards (e.g., ISO/IEC 29115 and ISO/IEC 24760) and reference implementations of remote attestation protocols that enable applications to assess the identity and integrity of devices.

4.2.3 Embedded Operating Systems Utilizing Hardware Security Features

Hardware designed for embedded systems is nowadays being extended with special hardware security features that enable the separation of the execution of untrusted code running in a "normal world" execution context from the execution of trusted code running in a "secure world" execution context. A number of new embedded operating systems have recently appeared but only few exploit hardware security features to their full extend. While some embedded operating system projects are truly open source, others are driven by vendors promoting specific hardware designs. As embedded hardware becomes increasingly powerful, it will be useful to converge on a common embedded software framework that supports a larger number of embedded hardware designs. Hence, it is desirable to develop a common European open source embedded operating systems utilizing hardware security features from the ground up. Ideally, this builds on existing expertise with open source embedded operating system activities that are not controlled or driven by a single vendor.

Actions:

- Development of open-source embedded real-time operating systems fully exploiting hardware security features. The operating systems should provide suitable abstractions and not be bound to specific proprietary hardware solutions.

4.2.4 Microkernel Isolation and Virtualization Mechanisms

In industrial environments and modern vehicles, the number of embedded control units is steadily increasing and reaching a point where consolidation is desirable since having separate embedded control units for each function is expensive and not scalable. Virtualization systems based on microkernel architectures start to become feasible and affordable for virtualizing embedded control units. However, more research needs to be done in order to achieve the level of isolation required for virtualizing safety-critical functions. In addition, functions need to be integrated that can continuously measure the integrity and separation that is being achieved.

Actions:

- Development of light-weight virtualization mechanisms for embedded devices that provide isolation and resource control satisfying the requirements for virtualizing safety-critical functions.

4.2.5 Open-source Secure Processor and Hardware Designs

Critical infrastructures require trust in all software and hardware components. The availability of well maintained open-source software has enabled the software industry to build software, including the software necessary to build software, from scratch using open-source components. On the hardware side, the industry typically relies on closed hardware designs and it has very limited tools at hand to verify whether a given piece of hardware is free from hidden functions or possible backdoors.

There is a movement towards open hardware designs. A prominent example at the processor level is the RISC-V project, providing an open-source CPU instruction set architecture enabling everybody to create RISC-V processors. Developing security extensions for RISC-V and hardware designs based on RISC-V technology will enable the industry to obtain hardware components from a variety of hardware components vendors, providing eventually the same control over the hardware components that is already possible on the software side.

Actions:

- Create an eco-system of open-source hardware designs enabling vendors to fully control the production of hardware components used in products controlling critical infrastructures.

4.2.6 Postquantum Cryptography on Constrained Devices

As quantum computers are evolving to a computational reality in the next few years, modern cryptography solutions (especially public key cryptography) need to be reinvented so as to avoid quantum processor-based cryptanalysis that can lead to full disclosure of secrets in reasonable amount of time. During the past years, the cryptography research community has invested time and effort to design and promote postquantum cryptography schemes that withstand quantum cryptanalytic attacks. NIST has launched a competition to award a standardized postquantum cryptography solution for key encapsulation mechanisms as well as digital signatures.

The European research community has a prominent role in this process with several schemes reaching the final competition round. The competition will be concluded in the upcoming years and the winner schemes will be broadly adopted by the security community. However, when such schemes are transferred to the Internet of Things environment and especially in resource constrained end nodes, there are several implementation aspects that need to be taken into account that

are not originally included in the postquantum cryptography algorithm definition. The relatively big cryptography keys used by the schemes as well as the computational complexity of those schemes may drain the resources of the existing IoT end node devices. The devices themselves may be deployed in a "hostile" environment where they may be attacked using side channel attacks. Furthermore, protocols for the IoT domain, like CoAP, do not yet take into account post-quantum cryptographic solutions and further adaptation at the protocol level should be made (e.g., on TLS or DTLS).

Actions:

- Adapt post-quantum cryptography solutions to the IoT and Industrial IoT environments so that they become deployable on resource constrained devices.
- Adapt existing IoT protocols to support post-quantum cryptography ciphers for key establishment and digital signatures.
- Investigate how to protect implementations of post-quantum cryptography schemes against side channel attacks including high order side channel attacks.
- Research novel lightweight post-quantum cryptography schemes that match the non-functional requirements of IoT end nodes and cyberphysical systems employed in the IoT/IIoT paradigm.

4.3 Impact of COVID-19 on Research Activities

The impact of COVID-19 on the research work in task T1.1 has been relatively minor. Since all planned physical meetings had to be canceled, a decision was made to start bi-weekly online meetings, which are documented on the internal Confluence project workspace. Initially, time was spent getting to get to know each others work in order to stimulate research collaborations. After the summer break, we started to invite external experts to give short introductions to specific technologies under standardization or to review the state of the art in research areas that are of interest of researchers active in T1.1.

- Henk Birkholz (Fraunhofer SIT, Germany) gave a presentation on October 29th about the IETF's work to standardize remote attestation procedures. His presentation was titled "Assessing Trustworthiness via Remote Attestation Procedures". Henk Birkholz is a co-author of several working group documents and one of the experts in this field.
- On November 26th, Zaid Hameed (Imperial College London, UK) provided on overview of the adversarial machine learning research done at Imperial College London. Zaid Hameed is a member of the Resilient Information Systems Security research group lead by Emil Lupu, who is also involved in the CONCORDIA project.

While the online meetings helped to create and maintain a community spirit, they do not fully replace face-to-face meetings, where new ideas are often born during social gatherings and spontaneous brainstorming sessions, which are difficult to replicated in scheduled online meetings. Hence, COVID-19 likely has a certain impact, in particular concerning the creation of novel ideas and research directions, but the originally planned research seems to be less affected by the pandemic.

In the next year, we plan to continue with regular online meetings, but we are considering to move to a monthly schedule. The invited presentations were perceived as useful and we plan to identify additional topics and presenters that can provide information of high relevance for researchers active in task T1.1.

4.4 Summaries of Research Activities

The following sections briefly describe research activities carried out during the year 2020.

4.4.1 Hardware Security for Embedded Systems / IoT Devices

Contact: Odysseas Koufopavlou (UP), Apostolos P. Fournaris (ISI)

Following the research directives that were defined in year one, University of Patras (UP) and the Industrial Systems Institute (ISI) of the Research Center ATHENA have been collaborating in order to expand the basis of this hardware cybersecurity research effort. During 2020, the outcome of this work is mainly focused on the design of embedded system hardware security sensors in the form of hardware security tokens, the possible adaptation of side channel attacks on postquantum schemes as well as the exploration of machine learning techniques in public key cryptographic implementations.

- A significant concern for the candidate schemes of the NIST postquantum cryptography standardization project is the protection they support against side-channel attacks. One of these candidate schemes currently in the NIST standardization race is the Dilithium signature scheme. This postquantum signature solution has been analyzed for side channel attack resistance especially against timing attacks. Expanding our attention on other types of side-channel analysis, our work is focused on correlation based differential side channel attacks on the polynomial multiplication operation of Dilithium digital signature generation [48]. In the paper, we describe how a Correlation Power Attack should be adapted for the Dilithium signature generation and describe the attack process to be followed. We determine the conditions to be followed in order for such an attack to be feasible, (isolation of polynomial coefficient multiplication inpower traces) and we create a power trace profiling paradigm for the Dilithium signature scheme executed in embedded systems to showcase that the conditions can be met in practice. Expanding the methodology of recent works that mainly use simulations for power trace

collection, in our paper, power trace capturing and profiling analysis of the signature generation process was successfully done on a noisy commercial off-the-shelf ARM Cortex-M4 embedded system.

- Anomaly detection systems (ADS), as part of a security information and event management (SIEM) system, are cybersecurity tools for identifying potential threats inside an information technology system. They are widely used in critical infrastructure (CI) systems for protection against attacks that can cause severe problems to public security and welfare. ADS collect information from various kinds of sources and correlate them to identify anomaly events. Such sources can be devices and software sensors, which inside a CI context (factories, power plants, remote locations) are placed in open areas and left unattended. These devices are vulnerable to tampering and malicious manipulation, which may then lead an ADS or SIEM system to ignore or falsely alert of possible cybersecurity problems. In our work, we developed strategies to mitigate the above problem using hardware means in order to enhance trust on ADS sensors [49]. Furthermore we propose a hardware/software based approach for legacy CI devices that can act as an ADS sensor or a tool for ensuring software ADS sensor data are not tampered.

- Deep learning-based side channel attacks are burgeoning due to their better efficiency and performance, suppressing the traditional side-channel analysis. To launch the successful attack on a particular public key cryptographic (PKC) algorithm, a large number of samples per trace might need to be acquired to capture all the minor useful details from the leakage information, which increases the number of features per instance. The decreased instance-feature ratio increases the computational complexity of the deep learning-based attacks, limiting the attack efficiency. Moreover, data class imbalance can be a hindrance in accurate model training, leading to an accuracy paradox. We propose an efficient Convolutional Neural Network (CNN) based approach in which the dimensionality of the large leakage dataset is reduced, and then the data is processed using the proposed CNN based model [97]. In the proposed model, the optimal number of convolutional blocks is used to build powerful features extractors within the cost limit. We have also analyzed and presented the impact of using the Synthetic Minority Oversampling Technique (SMOTE) on the proposed model performance. We propose that a data-balancing step should be mandatory for analysis in the side channel attack scenario. We have also provided a performance-based comparative analysis between proposed and existing deep learning models for unprotected and protected Elliptic Curve Cryptography (ECC) Montgomery Power ladder implementations. The reduced network complexity, together with an improved attack efficiency, promote the proposed approach to be effectively used for side-channel attacks.

This work constitutes a continuation of year one efforts on hardware cybersecurity and secure implementations, which are a close match of the activities of T1.1 Device-Centric Security, as well as the authentication use case regarding the T2.5 pilot on mobile communications. The upcoming research roadmap consists of a plan to upgrade the hardware security token platform to a more robust, versatile and lightweight solution, that can potentially become a host for possible implementations of postquantum schemes.

4.4.2 Secure Architecture for CyberPhysical System of Systems

Contact: Apostolos P. Fournaris, Aris Lalos (ISI)

A holistic, decentralized and cognitive design and operation approach has been described in [51] that supports CyberPhysical System of Systems (CPSoS) autonomic (without human intervention) behavior, making such systems aware of their physical and cyber environment and reacting to it accordingly so that they constantly match their intended purpose. We propose using a model based design approach to describe a CPSoS in a holistic and abstract way and to allocate computational power/resources to the CPS end devices of the System by determining and generating autonomously what cyber-physical processes will be handled by a device's each heterogeneous component (processor cores, GPUs, FPGA fabric) and software components (software stacks). The proposed solution uses this methodology to strengthen reliability, fault tolerance and security at system level but also to supports CPS designs that work in a decentralized way, collaboratively, in an equilibrium, by sharing tasks and data with minimal central intervention. Also, the proposed system supports the interaction of the CPSoS with their human users/operators through extended reality modules (AR glasses, haptics interfaces) to increase human situational awareness but also to include human behavior in the CPSoS design and operation phase. The proposal key points are highlighted in this paper and their usage in an automotive use case that involves connected cars is presented.

4.4.3 Device-centric and Attributes-based Authentication

Contact: Kostantinos Papadamou, Nikos Salamanos, Michael Sirivianos (CUT)

In the context of task T1.1, CUT is working on the implementation of a passwordless authentication solution that combines state-of-art technologies and protocols for killing the password and preserving privacy with device-centric and attribute-based authentication. More precisely, we leverage usable/strong device-centric authentication methods such as FIDO2[28] and Idemix[29] and we aim to provide a secure comprehensive authentication framework for users, as well as for IoT devices for

[28] https://fidoalliance.org/fido2/
[29] https://idemix.wordpress.com/

verifying their identity to remote services. More precisely, the developed solution will enable IoT devices to verify their footprint to a given server (IoT Controller), as well as to users, who have the privilege to access specific IoT devices to authenticate with them securely using strong authenticators such as fingerprint, face recognition, etc. Furthermore, with the developed solution we will be able to eliminate attacks like large denial of service attacks, as well as other types of attacks since IoT devices will have to verify their footprint/identity to a given server.

4.4.4 Architectural Support for Instruction Set Randomization

Contact: George Christou, Sotiris Ioannidis (FORTH)

Instruction Set Randomization (ISR) is able to protect against remote code injection attacks by randomizing the instruction set of each process. Thereby, even if an attacker succeeds to inject code, it will fail to execute on the randomized processor. The majority of existing ISR implementations is based on emulators and binary instrumentation tools, which unfortunately (i) incur significant runtime performance overheads, (ii) limit the ease of deployment, (iii) cannot protect the underlying operating system kernel, and (iv) are vulnerable to evasion attempts that bypass the ISR protection itself.

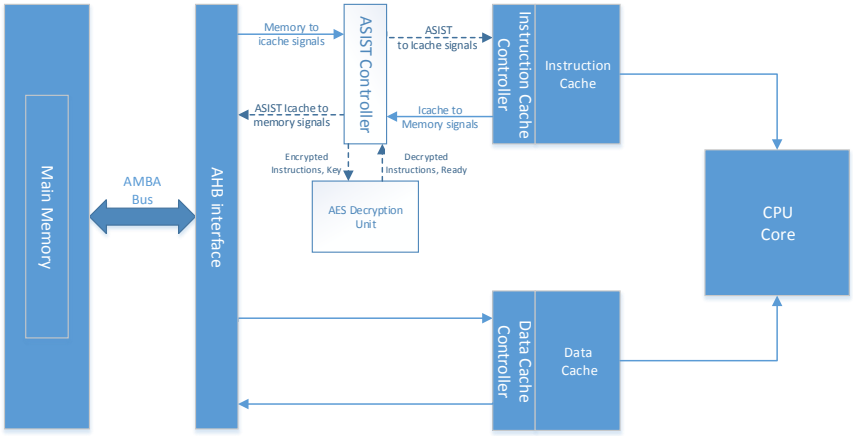

Figure 3: Hardware support for runtime instruction decryption

To address these issues, we implemented ASIST (Architectural Support for Instruction Set Randomization), as shown in the Figure 3, as a hardware/software scheme supporting ISR on top of an unmodified Instruction Set Architecture (ISA). We advocate that hardware support for ISR is essential to guard against user- and kernel-level code injection attacks, while the performance penalty incurred is within acceptable margins. To randomize the code instructions, we implement three different encryption algorithms, tailored for different needs and usages: (i) XOR, (ii) Transposition and (iii) AES. By adding the AES cipher, we can guarantee that an encryption key cannot be derived even if the attacker has access to both the

plaintext and the ciphertext (e.g., due to a memory leak). In addition, it can hinder any gadget discovery (which are a pivotal step of code reuse attacks) that are based on code pointer leaks in order to bypass address space layout randomization (ASLR) and the exploitation of memory disclosure vulnerabilities to map the text segment of a process. In both cases instructions will be encrypted with a strong cryptography scheme that prevents any form of cryptanalysis or bruteforce attacks.

4.4.5 Security Assessment of TrustZone-M based Software

Contact: Antonio Ken Iannillo (SnT)

Trusted hardware technologies are commonly used as anti-tamper technologies to make the software more resistant to attacks and protect critical program elements. It is generally more challenging to attack trusted hardware than a software-only protection scheme successfully. With the advent of the Internet of Things (IoT), embedded computers interact with the physical world or other software entities with minimal or no human input. ARM Holding, which already owns the largest share of mobile and embedded markets (60%), has further extended TrustZone-support for the tiniest low-end devices. ARM designed a hardware security extension from the ground up, instead of reusing it from application processors, for micro-controllers with the name of TrustZone Technology for Cortex-M profile or TrustZone-M.

While implementing the proposal from last year [74], we faced the challenges of interacting with such a novel technology. The reference secure software, namely Trusted Firmware M (TF-M), is still in its infancy. However, we detected that software designers are ignoring the issues TZ-M's software has. In particular, we are focusing on closing the semantic gap between the secure and non-secure world, exploiting the unique features of TZ-M. Indeed, the two worlds are completely isolated and unaware of each other's internal mechanisms. The secure world should have some insights of the non-secure internal protection mechanisms. Otherwise, the upcoming devices and their security services will be vulnerable to confused deputy attacks (e.g., a non-secure client access another non-secure client's private data by tricking the secure software). The implementation of our solution is ongoing, expecting a publication in 2021.

This research is relevant for the e-health pilot (T2.4) and the unmanned aerial vehicle pilot (T2.5). In the first case, the e-health pilot, particularly eesy-innovation, uses ARM processors for the connected medical devices. This pilot will eventually integrate TrustZone-M enabled devices, and this research may guarantee their security. In the second case, Airbus leads the implementation of a use case for UAVs that consists of authentication and authorizations between devices. The security of such communications can be guaranteed by ARM TrustZone-M devices and assessed by this research's outputs.

4.4.6 Trusted Execution Environment for IoT

Contact: Anum Kurshid (RISE)

The security of IoT devices is an ever-increasing concern with the expansion of the IoT network and the diversity of IoT devices incorporated in the underlying infrastructure. The IoT market and landscape constitutes vastly of IoT devices based on ARM processors (Cortex-A and Cortex-M family). RISE in collaboration with Jacobs University has an ongoing study regarding the vulnerabilities affecting ARM-based IoT devices. This study aims to paint a better picture of the vulnerability landscape in the IoT domain. This collaborative survey will also discuss the vulnerabilities and mitigation possibilities of IoT devices supporting TrustZone features.

Another effort under task T1.1 and T1.3, is exploring of Trusted Execution Environments (TEEs) as a mechanism to ensure device security. The newly introduced TrustZone-M technology brings the ability to establish TEEs in IoT nodes with limited resources and computational capabilities. Our research on TrustZone-M began last year with the identification of a lack of secure communication channel between the TrustZone worlds (normal and secure world). We designed and implemented a solution to (i) establish secure communication between the normal and secure worlds of a TrustZone-M enabled device and (ii) verify the authenticity of the normal world software requesting secure communication. Our research findings and the results of a systematic evaluation of the implemented prototype are formulated into a research paper "ShieLD: Shielding Cross-zone Communication within Limited-resourced IoT Devices running Vulnerable Software Stack", which is under submission for the IEEE Transactions on Dependable and Secure Computing.

4.4.7 Detection and Mitigation of Code-reuse Attacks

Contact: Abhilash Hota (JUB)

With a growing code base, embedded devices are increasingly vulnerable to code-reuse attacks in which code segments, called gadgets, from authorized software on these devices are reused for malicious activities. The work here aims to study and develop approaches to detect and mitigate code-reuse attacks [2, 121, 20] on ARM Cortex-M devices. The reliability of any models developed will depend on the reliability of the underlying data collection. ARM implements an isolated trusted execution environment using a security extension called TrustZone[9][110]. The use of ARM TrustZone for reliable instrumentation is to be studied and an appropriate instrumentation mechanism developed. The use of machine learning techniques is to be examined, for tasks like gadget and control-flow abnormality detection. Machine learning, and specifically neural networks, can require a lot of computational resources which are not available on the target platform. Hence, model compression techniques are to be used to develop lighter inference models that can run

on such resource-constrained devices without a substantial performance overhead. The initial literature review and research goals were presented at the CONCORDIA WP1 Early PhD workshop.

The work done so far has focused on surveying the state-of-the-art of attacks and mitigation strategies currently available on the Cortex-M platform. Approaches towards making lighter machine learning models have been surveyed. Currently, we are working on developing the binary instrumentation framework used to rewrite binaries for monitoring control flow, specifically any calls from non-secure software into the secure world.

Under a collaboration with RISE, vulnerabilities targeting ARM Cortex-A and Cortex-M processors are being surveyed. This study covers hardware and software attacks including fault attacks, side channel attacks, memory corruption attacks, call interception and just-in-time spraying.

4.4.8 Enclave Assisted Snapshot-based Kernel Integrity Monitor

Contact: Dimitris Deyannis, Sotiris Ioannidis (FORTH)

The integrity of operating system kernels is of paramount importance in order to ensure the secure operation of user-level processes and services as well as the benign behavior of the entire system. Attackers aim to exploit a system's kernel since compromising it provides more flexibility for malicious operations compared to compromising a user-level process. Acquiring access to the operating system kernel enables malicious parties to manipulate process execution, control the file system and the peripheral devices and obtain security and privacy-critical data. One of the most effective countermeasures against are kernel integrity monitors, implemented in software (often assisted by a hypervisor) or external hardware, aiming to detect threats by scanning the kernel's state. However, modern rootkits are able to hide their presence and prevent detection from such mechanisms either by identifying and disabling the monitors or by performing transient attacks.

Our efforts were focused on the implementation of SGX-Mon, an external kernel integrity monitor that verifies the operating system's kernel integrity using a very small trusted computing base while it does not require any operating system modifications or external hardware. In Fig. 4 we can see the architecture overview.

SGX-Mon is a snapshot-based monitor, residing in user space, and utilizes the trusted execution environment offered by Intel SGX enclaves in order to avoid detection from rootkits and prevent attackers from tampering its execution and operation-critical data. Our system is able to perform scanning, analysis and verification of arbitrary kernel memory pages and memory regions and to ensure their integrity. The monitored locations can be specified by the user and can contain critical kernel code and data. In Fig. 5 we can see the mapping of operating system kernel memory to the address space of the integrity monitor.

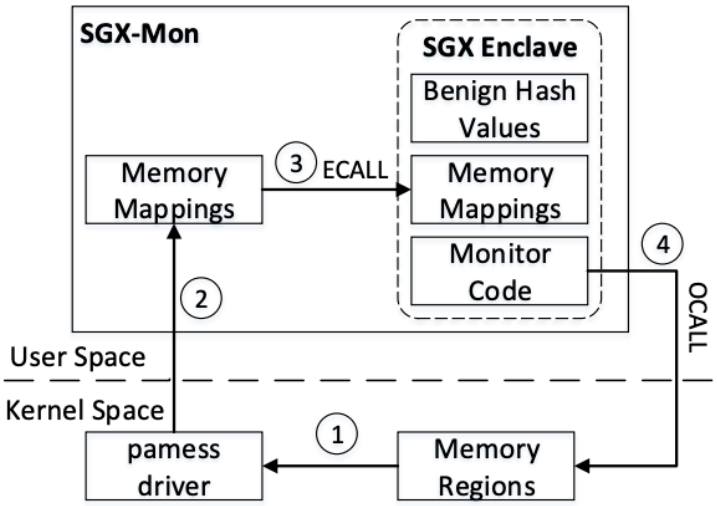

Figure 4: Architecture of the external kernel integrity monitor SGX-Mon

SGX-Mon scans the system periodically and compares the contents of critical memory regions against their known benign values. Our experimental results show that SGX-Mon is able to achieve 100% accuracy while scanning up to 6,000 distinct kernel memory locations [38].

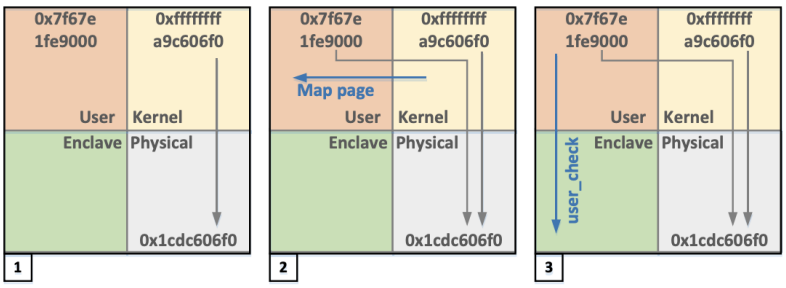

Figure 5: Mapping OS kernel memory to the address space of the integrity monitor

4.4.9 Privacy-preserving Malware Analysis in the Cloud

Contact: Dimitris Deyannis, Sotiris Ioannidis (FORTH)

While the number of connected devices is constantly growing, we observe an increased incident rate of cyber attacks that target user data. Typically, personal devices contain the most sensitive information regarding their users, so there is no doubt that they can be a very valuable target for adversaries. Typical defense

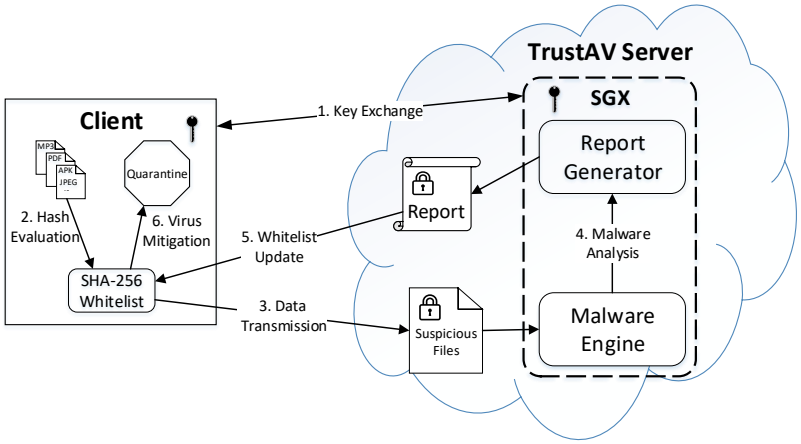

Figure 6: Architecture of the cloud-based malware detection solution TrustAV

solutions to safeguard user devices and data are based on malware analysis mechanisms. To amortize the processing and maintenance overheads, the outsourcing of network inspection mechanisms to the cloud has become very popular recently. However, the majority of such cloud-based applications usually offers limited privacy preserving guarantees for data processing in third-party environments.

Our efforts were focused on the implementation of TrustAV, a practical cloud-based malware detection solution destined for a plethora of device types. TrustAV is able to offload the processing of malware analysis to a remote server, where it is executed entirely inside, hardware supported, secure enclaves. By doing so, TrustAV is capable to shield the transfer and processing of user data even in untrusted environments with tolerable performance overheads, ensuring that private user data are never exposed to malicious entities or honest-but-curious providers. In the Figure 6 we can see the architecture overview.

TrustAV utilizes various techniques in order to overcome performance overheads introduced by the Intel SGX technology. It reduces the required enclave memory,a limiting factor for malware analysis executed in secure enclave environments, offering up to three times better performance [39].

4.4.10 Distributed Protected Execution System

Contact: Dimitris Deyannis, Sotiris Ioannidis (FORTH)

During this period we designed and implemented a distributed system that allows the parallel execution of JavaScript code on multiple remote nodes that support Intel SGX enclaves. Figure 7 we show the architecture of this system.

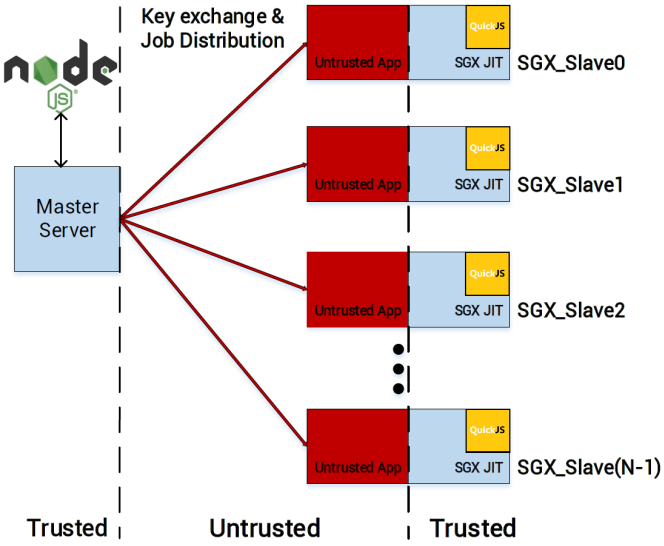

Figure 7: Architecture of parallel JavaScript code execution in SGX enclaves

Our system is composed of a set of libraries, residing in the Node.js framework, that allow developers to mark the desired JavaScript functions for secure parallel execution. Once the functions are tagged, our framework analyses them and via a custom scheduler initiates their parallel execution on a set of remote SGX-assisted nodes. These nodes execute the offloaded JavaScript functions using the QuickJS framework, which we have ported to run entirely within SGX enclaves. The process of offloading the JavaScript code is performed transparently to the developers which are only required to mark specific functions with our custom tags.

4.4.11 Machine Learning for IoT Device Identification

Contact: Asaf Shabtai, Yair Meidan (BGU)

At Ben-Gurion University (BGU), a variety of research efforts have been undertaken during the past year regarding device-centric security. Two of these research projects ([81] and [91]) address the challenge of IoT device (model) identification, another project [47] proposes a method for detecting adversarial inputs which might be used to bypass IoT identification systems, and another one [90] explores the connections between IoT device complexity, traffic predictability and attack detectability. Following is an overview of each of the above-mentioned research efforts, accompanied by the associated publication.

1. **IoT device identification using deep learning.** To address the risks posed to organizations by the widely adopted *bring your own device* (BYOD) policy, we developed a traffic-based method which enables white-listing of BYOD IoT devices [81]. To protect their networks, organizations must be able to

identify the IoT devices connected to their networks and, more specifically, to identify connected IoT devices that are not on the white-list (unknown devices). In contrast to previous work, our approach does not require that complex feature engineering be applied on the network traffic, since we represent the communication behavior of IoT devices using small images built from the IoT devices' network traffic payloads.

2. **Identification of vulnerable IoT models behind a home network NAT.** Telecommunication service providers (telcos) are exposed to cyber-attacks executed by compromised IoT devices connected to their customers' networks. Such attacks might have severe effects on the attack target, as well as the telcos themselves. To mitigate those risks, we developed [91] a machine learning-based method that can detect specific vulnerable IoT device models connected behind domestic network address translation (NAT), thereby identifying home networks that pose a risk to the telcos infrastructure and service availability. Our experimental results show that the flow-based method we propose is robust and can handle situations for which existing methods used to identify devices behind a NAT are unable to fully address, e.g., encrypted, non-TCP or non-DNS traffic.

3. **Detection of adversarial examples using SHAP signatures.** Although typically effective in many application domains, deep neural networks (DNNs) are vulnerable to adversarial perturbation attacks. In this research [47], we presented a novel detection method that uses Shapley Additive Explanations (SHAP) values computed for the internal layers of a DNN classifier to discriminate between normal and adversarial inputs. We evaluated our detector against adversarial examples generated by diverse state-of-the-art attacks and demonstrated its high detection accuracy and strong generalization ability to adversarial inputs generated with different attack methods. In the future, we plan to evaluate this detector in the context of IoT device identification, where adversaries might tamper with the network traffic of rogue IoT devices in order to bypass IoT identification and white-listing systems (e.g., [81] and [91] which we published this year).

4. **Mining and monitoring darknet traffic for threat detection.** Trillions of network packets are sent over the Internet to destinations which do not exist. This "darknet" traffic captures the activity of botnets and other malicious campaigns aiming to discover and compromise devices around the world. We have developed a framework and algorithm for mining darknet traffic called DANTE [28]. DANTE learns the meaning of targeted network ports by applying Word2Vec to observed port sequences. To detect recurring behaviors and new emerging threats, DANTE uses a novel and incremental time-series cluster tracking algorithm on the observed sequences.

5. **Quantification of IoT-related attack detectability.** IoT devices are known to be vulnerable to various cyber-attacks, such as data exfiltration and the

execution of flooding attacks as part of a DDoS attack. When it comes to detecting such attacks using network traffic analysis, it has been shown that some attack scenarios are not always equally easy to detect if they involve different IoT models. That is, when targeted at some IoT models, a given attack can be detected rather accurately, while when targeted at others the same attack may result in too many false alarms. In this research [90], we attempted to explain this variability of IoT attack detectability and devise a risk assessment method capable of addressing a key question: *how easy is it for an anomaly-based network intrusion detection system to detect a given cyber-attack involving a specific IoT model?* While addressing this question we *(a)* investigated the predictability of IoT network traffic, *(b)* presented a novel taxonomy for IoT attack detection which also encapsulates traffic predictability aspects, *(c)* proposed an expert-based attack detectability estimation method which uses this taxonomy, and *(d)* empirically evaluated our method while comparing it with a data-driven method.

4.4.12 IoT Identification with A Physically Unclonable Function

Contact: Sina Rafati Niya, Burkhard Stiller (UZH)

The integration of Internet-of-Things (IoT) and Blockchains (BC) for trusted and decentralized approaches enabled modern use cases, such as supply chain tracing, smart cities, and IoT data marketplaces. For these it is essential to identify reliably IoT devices, since the producer-consumer trust is not guaranteed by a Trusted Third Party (TTP). Therefore, the work designed and implemented by the University of Zurich (UZH) proposes a Know Your IoT device platform (KYoT), which enables the self-sovereign identification of IoT devices on the Ethereum BC. KYoT permits manufacturers and device owners to register and verify IoT devices in a self-sovereign fashion, while data storage security is ensured. KYoT deploys an SRAM-based (Static Random Access Memory) Physically Unclonable Function (PUF), which takes advantage of the manufacturing variability of devices' SRAM chips to derive a unique identifying key for each IoT device. Due to the SRAM's bias toward 0 or 1, KYoT employs a fuzzy extractor to generate a unique key from noisy SRAM data, as well as a helper, which is needed in order to reproduce the unique key in future verification processes. KYoT uses 32 Byte input strings to optimize storage costs. Hence, the number of allowed noisy bits is set to 8, sufficiently enabling the reproduction of keys reliably in all tests. The self-sovereign identification mechanism introduced is based on the ERC 734 and ERC 735 Ethereum identity standards. ERC 734 describes standard functions for uniquely identifiable proxy SCs that can be used by other accounts or SCs. These SCs describe "anything", such as groups of individuals or devices, and act as an identity proxy on the BC. The described identity SC has a key-storage controlling the degree to which other parties can interact with the contract. E.g., upon creation of the contract, a management key is added to the key-storage based on the creator's account ad-

dress to ensure that certain functionality can only be executed by the SC creator. The contract also features an execute function to run arbitrary contract calls, which acts as a proxy for whatever instance it is representing. ERC 735 extends these functions to add and remove claims, which are hashed and signed by an identity contract of a trusted third party. In order to issue claims, this contract must first add a claim key to its key-storage, as it will be used to sign the claim and will be later checked by anyone wishing to check, if the claim is valid. The design, implementation, and evaluation of KYoT has been presented at the LCN 2020 conference [116] as full paper in the Ph.D. track.

4.4.13 Security Assurance for IoT Systems

Contact: Claudio A. Ardagna, Marco Anisetti, Lara Mauri (UMIL)

The existence of billions of cheap and resource-constrained devices connected to the Internet introduces fundamental risks that can threaten users' life and personal sphere. A wealth of services in different domains, such as smart vehicles, smart buildings, e-health, are distributed on the basis of data collected by devices. In this context, assurance evaluation is fundamental to guarantee the correct behavior of the whole system and its devices. The term assurance means, in a wider sense, the technical judgment that a service, process, or device satisfies some properties. The implicit assumption is that data have a sufficient level of trustworthiness to create information, and in turn knowledge and wisdom. This assumption is, however, not sound when a plethora of devices are used to collect data, and might lead to scenarios where wrong evidence results in wrong decisions and, in turn, untrusted services/applications. It is likely that without an open, protocol-neutral baseline solution for IoT assurance, fundamental risks will create further exploitation opportunities. Research on IoT-based systems assurance was recently started and mainly focused on defining new assurance architectures for IoT.

In the first year of the project we set up the stage for a general approach of assurance evaluation in IoT environments. In the second year, we worked around three main topics. First, we continued our work that focused on providing a trustworthy IoT environment, where the automation and adaptation processes are based on trustworthy data [14]. The idea described in this paper has been extended in the second year of the project by providing *i)* an architecture for trustworthy data collection based on blockchain and smart contracts, *ii)* a methodology for the assessment of the trustworthiness of collected data based on syntactic and semantics rules, *iii)* different processes for collecting, verifying, and storing trustworthy data, which balance trustworthiness and privacy. We further evaluated the performance of our approach and its quality in a simulated scenario considering data collected from smart homes and in a real-world scenario considering roaming data collected from a switching telco infrastructure. The extended paper has been accepted for publication in ACM Transactions on CyberPhysical Systems and is currently in the production phase.

Then, we activated a parallel line of research on specific application verticals. More in details we focused on unoccupied aerial systems (UAS) in the context of a shared urban airspace where cooperation among vehicles with different levels of trustworthiness is needed [11]. In such a context, assurance evaluation of the vehicles will help in evaluating the trustworthiness and thus on improving the reliability of cooperation between them. To support efficient and safe collaboration among UASs, the following steps and challenges need to be taken into account: *i)* detecting the incoming vehicles, *ii)* establishing a secure communication channel, *iii)* determining the level of interaction. We focused our attention on the last point and initially discussed gaps and challenges, and a possible approach to trustworthy coalitions and collaborations based on assurance evaluation. We first detailed the reference scenario starting from the vehicle capabilities and collaboration intensive use cases, we then described how the vehicles can safely cooperate in such context. Starting from this scenario we elicited requirements describing challenges and actions to take to cope with the challenges. More specifically we identified: *i)* the need to estimate a trust indicator per vehicle, *ii)* the need to adapt interactions between vehicles based on the trust indicators, *iii)* the need to control the information sharing, *iv)* the need of continuous monitoring and assurance. Concluding, we identified a roadmap on UAS security that will drive the research also in the area of assurance for this specific sector. The outcome of this work is important for the pilot "Security of Unmanned Aerial Systems (UAS)", where the security of unmanned aerial systems must be protected and passes from the quality of collected data used, for instance, for drone authentication.

Finally, we considered a topic that has been mainly carried out in Task 1.1, and is also relevant for activities in Task 1.4, aiming to evaluate the assurance of IT systems including IoT devices in the cloud or on premises. In [87], we introduced the notion of Reciprocally Useful Work (RUW), a novel update mechanism for Distributed Ledgers (DLs) where any agent wishing to add a block to the ledger must first perform an activity that will improve the utility for the DL-supported application of some other agent's block. We demonstrated that such RUW-supported trust in training data can be used to alleviate the problem of poisoning attacks to Machine Learning (ML) models. When applied to DLs storing training data for ML models, RUW can play the role of a direct compensation of the potential disruption, which is measurable in term of the performance of the ML model trained on the DL content. Thus, our ML-oriented distributed consensus mechanism can support ML models' training set selection. A proper assurance solution with advanced chain of trust is important for all pilots providing functionalities for evaluating the security status of a system and its level of compliance to predefined rules. In particular, together with TELENOR and the pilot "Threat Intelligence for the Telco Sector", we are progressing in investigating the possibility of applying the proposed assurance techniques to 5G communications.

4.4.14 RESTCONF Support for OpenWrt Devices

Contact: Jürgen Schönwälder (JUB)

OpenWRT is an embedded Linux operating system, which is popular on customer premises equipment and on access points in wireless home networks. It runs on many different and affordable types of hardware. OpenWRT is also used by many community and research projects such as the freifunk mesh networking project, distributed network measurement projects, bufferbloat measurement and experimentation tools, or the multi-path TCP project. Devices running OpenWRT can typically be accessed and configured via SSH and an embedded web-based user and configuration interface. OpenWRT also supports traditional network monitoring protocols, such as NetFlow, IPFIX, or SNMP.

Modern device management interfaces are often defined using the YANG data modeling language [19]. Data model driven protocols such as NETCONF [46] or RESTCONF [18] are then used for device configuration and monitoring. Data models exist for many network functions. For IoT devices, the manufacturer usage descriptions [85] (policies provided by device vendors and enforced on network access devices) are gaining popularity and data models for remote device integrity verification are under preparation.

In order to experiment with YANG-defined interfaces on embedded devices running OpenWRT, a lightweight RESTCONF implementation called ORC has been implemented from scratch targeting OpenWRT devices. ORC has been designed to use as little resources as possible when no RESTCONF interactions take place and to integrate well into the typical OpenWRT software infrastructure. A paper describing the design of ORC has been presented in the experience paper track of the IEEE/IFIP NOMS 2020 conference [60], where it received the best experience paper award.

4.4.15 Blockchain-based Industrial Internet of Things Systems

Contact: Sina Rafati Niya, Eryk Schiller, Burkhard Stiller (UZH)

In Industry 4.0 (I4), the Industrial Internet of Things (I2oT) data streams are prone to significant data manipulation risks. The integration of Blockchains (BC) with I2oT may become a solution preventing from this problem. UZH provided two papers in this direction [103, 126]. The paper presented at IEEE/IFIP NOMS 2020 provided a blockchain-agnostic Blockchain I2oT (BI2oT) architecture called BIIT that allows developing a broad range of BC applications fully integrating Internet of Things (IoT). The follow up of this research was published at IEEE LCN 2020. The mechanisms introduced in BIIT aim at solutions that provide extensive data reliability, limit the computational overhead, and enhance energy efficiency. BIIT is evaluated through real-world experimentation using the Atmel AVR device family with Long Range (LoRa) and LTE-M communication technologies [103] as well as

the Texas Instruments MSP430 device family with IEEE 802.15.4 communication technologies [126].

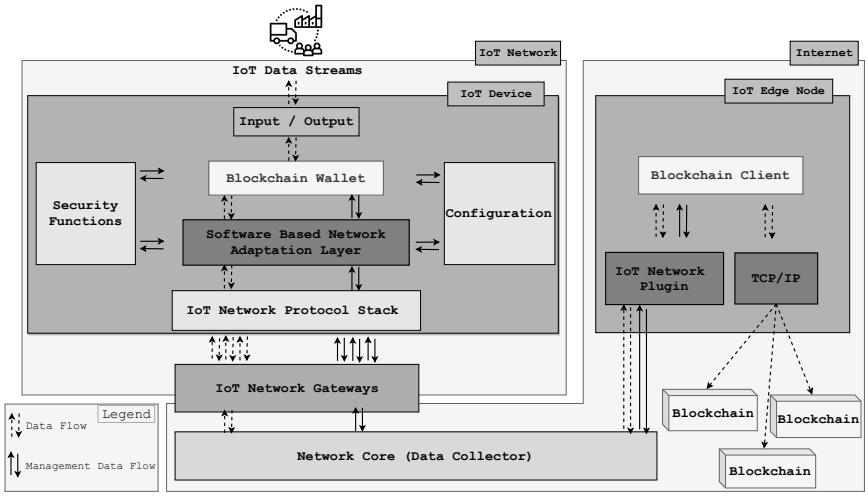

Figure 8: BIIT: Blockchain-IoT Compliance Architecture for Industry 4.0

The BIIT architecture shown in Fig. 8 defines the following components on IoT devices and uses them to offer the BI2oT integrated management:

- IoT Devices are end devices used for data collection that communicate with other systems components through the IoT network infrastructure.
- The BC Wallet is a software component placed on an IoT device that contains device specific credentials such as BC address, private key, destination addresses, balance, transaction (TX) counter, etc. This information is essential to provide data authenticity, integrity, and trust. Using the Configuration engine, the BC wallet configures the lower layer components, *i.e.,* Network Adaptation Layer (NAL), to produce appropriate TXs optimized for the underlying network technology. Moreover, the BC wallet issues data packets that shall be converted through the NAL into fully fledged TXs on the southbound interfaces. The data packets may contain information on the the industrial process monitored.
- Security Functions: Similar to the Security Service Provider defined in ZigBee[30], a generic BIoT architecture has to consider a security engine, *i.e.,* Application Programming Interfce (API), producing a large number of hash functions (*e.g.,* SHA-256) and Elliptic Curve Digital Signature Algorithm (ECDSA) signature types (*e.g.,* Ed25519[31]), which guarantee the compatibility with a large spectrum of BC protocols.

[30] https://zigbeealliance.org/
[31] https://ed25519.cr.yp.to/

- Configuration: The configuration engine manages the configuration of the wallet and Network Adaptation Layer (NAL). The Wallet is configured by the Enterprise End User (EEU) using the management plane. In turn, the configuration of the NAL is requested by the wallet, which demands the computation of appropriate hash functions (*e.g.,* SHA-256) or ECDSA signatures (*e.g.,* Ed25519). Furthermore, the Configuration engine provides the necessary keys (*e.g.,* Ed25519 private key) to the NAL and negotiates the data-plane message format used by the NAL on the southbound network interfaces.
- The Software Based Network Adaptation Layer (NAL) is a supporting component of the wallet; it receives data packets from the wallet on the data-plane, computes the required hash functions (*e.g.,* SHA-256) or ECDSA signatures (*e.g.,* Ed25519), assembles TXs, and sends them to the underlying network protocol stack. NAL optimizes the packet transmission towards the underlying network layer (*i.e.,* message fragmentation, Automatic Repeat reQuest (ARQ), TX aggregation) and provides TX compatibility with the targeted BC (*e.g.,* BTC, ETH). The adaptation is required as in many situations, the TXs cannot be sent directly by the network layer, *e.g.,* the regular TX size may already exceed the Maximum Transfer Unit (MTU) of the IoT network.
- The IoT Network Protocol Stack comprises the physical, data link, and networking layers enabling the communication between IoT nodes and IoT Gateways (GWs), *e.g.,* LoRa GW, and LTE Machine Type Communication (MTC) evolved Node B (eNB)).
- IoT Network Gateways relay data packets between the IoT network and other computing infrastructures, *e.g.,* Edge-Nodes and cloud infrastructures.
- The Network Core provides the data collection point, in which the IoT data is stored temporarily and can be accessed by other system components (*e.g.,* BC Client).
- Since BIIT tackles heavily constrained devices, IoT nodes are not envisioned to play the role of full BC clients. The IoT Edge Node functionality may be provided on less resource constrained devices such as a Raspberry PI (RPI)[32] helping the BC Wallets on IoT devices. In some cases, *e.g.,* in LoRa, the Edge Nodes derive the TXs from TX chunks sent by the IoT device as for example in LoRa, the regular TX size may exceed the LoRa MTU. It is worth noting that some Edge Nodes can also act as a gateway relying packets from the IoT network towards other computing infrastructures.

[32]https://www.raspberrypi.org/

- Typical full BC clients require significant amounts of resources (*e.g.,* storage at the order of GB in the case of BTC[33] or ETH[34]) and are not implementable on IoT devices. The BC Client residing on the Edge Node is an auxiliary element helping out IoT wallets to submit/verify TXs using the Edge Node resources.
- The IoT Network Plugin receives the IoT data stored in the network core, *e.g.,* using Message Queuing Telemetry Transport (MQTT) to access data, and provides the collected packets towards the BC Client using an API.
- The BC client communicates using the TCP/IP stack with other computing infrastructures (*e.g.,* BC miners).
- Any type of Blockchain (*e.g.,* BTC, ETH) can attached to the BIIT architecture working on physical or cloud infrastructures.

The developed architecture allows for easy management of IoT devices and data streams in the industrial setup, and provides high performance protocols for different IoT network access technologies. The architecture has been evaluated in various setups including different hardware architectures, *i.e.,* Atmel AVR and Texas Instruments MSP430, using several communication technologies such as LoRa, LTE-M, and IEEE 802.15.4.

4.4.16 SDR-based LoRa Cloud Radio Access Network

Contact: Eryk Schiller, Burkhard Stiller (UZH)

Long Range (LoRa) defines a popular modulation scheme based on the chirp spread spectrum technique. It is used in Low Power Wide Area Networks (LP-WANs) for IoT. Our work, published at WiMob'20 [127], designs, specifies, implements, and evaluates a Cloud Radio Access Network (CRAN) architecture for LoRa networks, while using Software Defined Radios (SDRs) to receive/send radio signals and Docker to virtualize the setup. A software modulator is developed to emit signals on the downlink targeting regular LoRa end-device receivers, such as Semtech SX1276 chips. Finally, the network, processing, and cost requirements of the C-RAN implemented are evaluated.

Fig. 9 displays the high level architecture with all modular components involved. It contains three computing entities, *i.e.,* Base Band Unit (BBU), Remote Radio Head (RRH), and Network Server (NS). The RRH has an SDR attached, *e.g.,* through the Universal Serial Bus (USB) port, to send and receive radio signals using the SDR radio chain.

Regular devices generating and receiving LoRa-compliant signals are this architecture's clients. An example Arduino device may emit radio waves using a reg-

[33] https://bitcoin.org/
[34] https://ethereum.org/

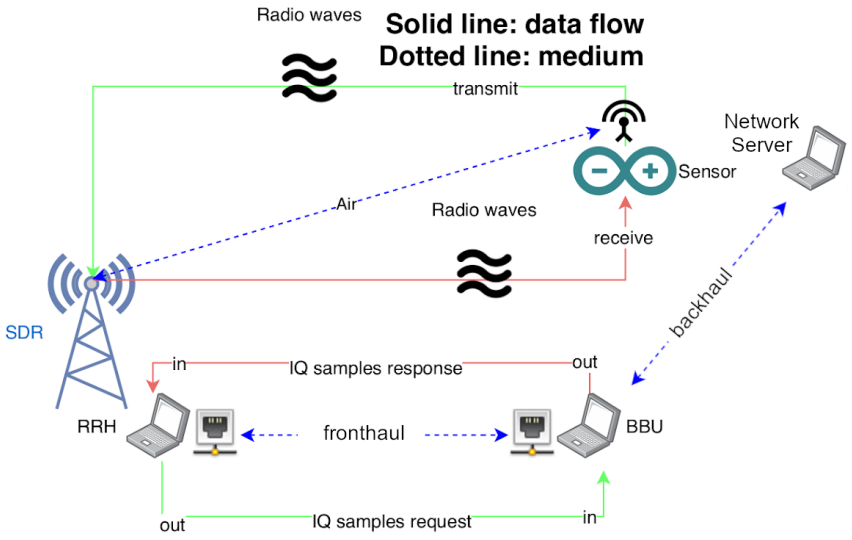

Figure 9: High Level C-RAN Architecture

ular SX1276/SX1278 radio transceiver compatible with the LoRa PHY manufactured by Semtech. It is worth noting that LoRa PHY is patented, and only a limited set of information about the chip design is known by the public. However, SX1276/SX1278 transceivers are Semtech-compliant, and, therefore, operate using waveforms compatible with all Semtech-based devices. As such, regular sensors (*e.g.,* Arduino-based) equipped with legitimate hardware, *e.g.,* SX1276, may benchmark the reverse-engineered software-based LoRa PHY implementation.

Radio waves coming from sensors on the uplink are picked up by the RRH, which samples analog signals and sends the corresponding 32-bit In-Phase and Quadrature Components (I/Q) sample stream over the fronthaul network to the BBU (*cf.* Fig. 9). The BBU entity runs a software-based LoRa modem as a Virtual Network Function (VNF). The virtual LoRa modem demodulates and decodes the I/Q sample stream. The message decoded is sent to the NS for processing on the upstream using the backhaul interface. In the case a response is requested, the NS originates a down-stream data packet towards a BBU through the backhaul interface. This response packet is encoded by the BBU and sent as resulting down-stream I/Q samples to the RRH on the fronthaul interface. The RRH provides the sample stream received toward the attached SDR, which in turn emits the analog radio signal through the outgoing radio chain so that the signal is transmitted back to the Arduino on the downlink. Finally, the Arduino device receives the LoRa waveform using the regular SX1276/SX1278 radio transceivers.

Moving toward a C-RAN (Cloud Radio Access Network) in LoRa (Long Range) is now proven feasible as the implementation and evaluation performed have shown. To make this happen, LoRa gateways' functionality had to be split into separate

RRH (Remote Radio Head) and BBU (Baseband Unit) components, where each can be containerized and run in separate, virtualized environments. However, this approach leads to a measurable amount of fronthaul load of approximately 2 MiB/s between the RRH and the BBU for the 125 kHz BW case as proven experimentally. Since SDRs (Software-Defined Radio) enable the reception and sending of radio signals., especially, the software modulator developed emits successfully signals on the downlink in support of regular LoRa end-device receivers. Nevertheless, such a C-RAN offers more flexibility in the long run against LoRa PHY (Physical Layer) hardware solutions, because amendments to the LoRa PHY can be provided in software, when required, without any infrastructural changes on a broader range. The evaluation of network, processing, and cost requirements of the new C-RAN implemented indicate besides a clear feasibility that communications are possible at a reasonable cost, including the technical efficiency in certain cases.

Thus, the three goals have been achieved successfully, the architecture and its specification, a detailed evaluation of a C-RAN for LoRa, and based on a reverse engineered LoRa PHY the extension of an existing uplink signal encoder to generate downlink signals being compliant with regular LoRa transceivers SX1276/1278.

4.4.17 MQTT-based Privacy Preferences Enforcement

Contact: Elena Ferrari, Pietro Colombo (UI)

We have extended the privacy preference model previously proposed in [23] by exploiting an attribute-based access control (ABAC) model to control the communication of IoT devices operating in a single MQTT environment [29]. The extension aims at regulating data exchange in scenarios where data originated from IoT devices span the boundaries of the organization where they have been originally collected. The proposed access control framework provides support for access control policies and user preferences that regulate data sharing across bridged MQTT environments [32]. The key contribution of [32] is the decentralized approach that allows enforcing access control policies and user preferences specified in each side of a pair of interconnected MQTT environments. User preferences can even be enforced in environments different from the ones where they have been originally specified. In [32] a new enforcement monitor that regulates the messages that can enter or leave an environment has been introduced, which operates along the environments bridge. The decentralized enforcement mechanism has been implemented relying on the joint work of this new monitor and of monitors deployed in each environment of a bridged pair. To the best of our knowledge, the approach presented in [32] is among the earliest edge-based access control approaches that allow regulating data sharing across interconnected IoT environments.

The resulting ABAC framework for MQTT environments has then been further enhanced with a dashboard of analysis services designed for security administrators [31]. As a matter of fact, due to the potentially numerous MQTT clients that op-

erate in an MQTT environment, and to the numerous access control policies and user preferences specified for this system, it is fundamental that security administrators could use tools to analyze the effects of policies and preferences on the exchanged messages, as well as the decisions taken by the enforcement monitor. This is particular crucial since in the considered ABAC framework authorizations are implicitly denoted through conditions over subject, object and environment attributes. The dashboard has been specifically designed to fulfill all aforementioned requirements, and allows administrators to perform real-time monitoring of the enforcement monitor activities. It provides services to analyze the message flow produced and consumed by any MQTT client, and the access control policies and user preferences that regulate the access to messages on given message topics, and the effects of the specified policies/preferences on the exchanged messages.

Afterwards, we have further extended this framework, to regulate information sharing in MQTT-based IoT environments during emergency situations. We propose a framework that allows security administrators to specify patterns of events that trigger emergencies (e.g., a fire, a fall, a heart attack), and access control policies that regulate data sharing during an emergency situation, as well as in case of no emergency. The approach relies on a complex event processing system to detect emergency conditions, and on a situation-aware enforcement monitor, which enforces policies applicable in the identified context. Early experimental evaluations show the efficiency of the proposed mechanisms.

4.4.18 Privacy Preservation in Industrial IoT

Contact: Apostolos P. Fournaris, Aris Lalos (ISI)

The Industrial Internet of Things (IIoT) is a key element of industry 4.0, bringing together modern sensor technology, fog and cloud computing platforms, and artificial intelligence to create smart, self-optimizing industrial equipment and facilities. Though, the scale and sensitivity degree of information continuously increases, giving rise to serious privacy concerns. We proposed in [83] an efficient privacy preservation technique that tracks the correlation of multivariate streams recorded in a network of IIoT devices. The time-varying data covariance matrix is used to add noise that cannot be easily removed by filtering, generating obfuscated measurements and, thus, preventing unauthorized access to the original data. To improve communication efficiency between connected IoT devices, we exploit inherent properties of the correlation matrices, and track the essential correlations from a small subset of correlation values. Extensive simulation studies using constrained IIoT devices validate the robustness, efficiency, and effectiveness of our approach.

5 Network-Centric Security (T1.2)

Task 1.2 of the CONCORDIA project relates to network-centric security. T1.2 has identified three broad areas of research within the network-centric security context: 1) making infrastructure and services (more) resilient against (Distributed) Denial-of-Service ([D]DoS) attacks; 2) encrypted network traffic analysis, particularly to the end of detecting network-based threats; and 3) using Software-Defined Networking (SDN) to improve network resilience and stability.

The intentions within T1.2 are to:

- investigate proactive, coordinated and distributed strategies to defend against network-based attacks, (D)DoS in particular;
- collect, share and analyze data on attacks and intrusion to the end of mitigation, attribution and other actionable intelligence;
- develop techniques to monitor encrypted traffic for security purposes; and
- investigate SDN as a means to form a trusted and resilient Internet for Europe.

5.1 Overview

In the following paragraphs, we will briefly touch upon research efforts undertaken in the T1.2 context in year 2. Each mention will be accompanied by a reference to a later, more detailed discussion (placed under Section 5.4). Before arriving at the more elaborate sections, we will outline how the undertaken efforts relate to the pilot projects of the CONCORDIA project. Where applicable, we also consider how continued efforts relate to their counterpart in year 1.

Bird's-Eye view on year 2 efforts
The domain name system (DNS) is core component of the Internet. Many networked services depend on the DNS. Various T1.2-related efforts have focused on the DNS for this reason, more so in year 2. Two directions of DNS-based related research were taken. First, making the DNS itself more resilient against attacks. And second, making the DNS less viable to bring about attacks, examples of which are DNS-based reflection and amplification attacks.

In Section 5.4.1 we describe work that identifies DNS configuration errors that may be leveraged by malicious actors to undermine the stability of the DNS, as well as to potentially intercept and hijack network traffic.

In terms of abuse of the DNS, we investigated security threats lingering in a specific type of DNS record: text resource records (`TXT`). In Section 5.4.2 we discuss work that uncovered various TXT-related issues that have security implications. Another form of abuse of the DNS relates to reflection and amplification attacks. DNS security extensions (DNSSEC) plays a prominent role here because of the

amplification potential brought about by the addition of cryptographic signatures. Section 5.4.3 describes efforts to bring quantum safe algorithms to DNS security, which reduce signature sizes and thus the potential for amplification attacks.

Various T1.2 studies put IP anycast in focus during the second year. IP anycast is a technology to make networked services more resilient against attacks. While various existing IP anycast deployments cover DNS infrastructure, anycast extends beyond any specific networked service.

In Section 5.4.4 we introduce Manycast^2, a novel methodology to assess whether Internet services use anycast. Armed with knowledge about anycast deployment, the resilience properties of services can be studied. Section 5.4.5 introduces another tool in the anycast ecosystem: BGP Anycast Tuner. Rather than finding existing IP anycast deployments, BGP Anycast Tuner can aid network operators bring about and bolster anycast deployments.

Threat intelligence and intrusion detection are essential to network-centric security. These topics continue to be an important part of T1.2-related research, equally so to year 1. Taking a proactive approach to detecting security threats is different from a, arguably more traditional, reactive approach. In Section 5.4.6, we report on a study that discusses the pitfalls of a proactive approach and suggest ways of how these can be overcome.

Monitoring network traffic is essential to the end of identifying network-based threats. In year 1, various efforts in the T1.2 context were started to lower the barriers to analyze encrypted network traffic. Several of these efforts have further evolved over year 2. Section 5.4.7 focuses on analyzing encrypted HTTP/2 traffic, so as to be able to monitor such traffic for signs of illegitimate use of services, in a privacy-preserving matter. Section 5.4.8 reports on flow-based measurement and analysis of TLS and QUIC protocols to provide situational awareness in the ever-evolving computer networks. Section 5.4.9 outlines HeadHunter, a fast, signature-based intrusion detection system that can even operate on encrypted network traffic by considering metadata extracted from packet headers. To the end of scaleability, HeadHunter in part can be hardware (GPU) accelerated.

Compromised IoT devices can pose a threat to the security and stability of networks and are regularly used to bring about attacks. Having the ability to fingerprint and identify IoT devices at the network edge is important to identify threats against network security. Section 5.4.10 reports on a federated learning approach that allows for IoT device fingerprinting.

Data is a key enabler for network-centric research. Acquiring and developing quality data sources is often a first step to detecting threats, studying resilience, and gaining actionable intelligence. In Section 5.4.11 we report on extensive efforts taken to acquire and developing (raw) data sources to methodologically study parts of the DDoS ecosystem. Specifically, we extensively study: 1) (D)DoS attacks; 2) attacked targets; and 3) two popular technologies to mitigate attacks, which has

also allowed us to identify common mistakes in deployment that have grave consequences for the user. Other data-related efforts within the T1.2 context concern the generation of high-quality encrypted traffic datasets. Section 5.4.12 reports on work that builds on successes from the first year, to synthesize datasets with non-trivial attacker behavior.

As we reported before for year 1, blockchain can be considered in a dual way when it concerns cybersecurity. First, as an attractive target for attackers that should be protected. Second, as a tool to build novel, collaborative defense mechanisms. Various T1.2 efforts have kept focus on blockchain in year 2 for this reason. Section 5.4.13 describes efforts to facilitate scalability for solutions built on top of blockchain by introducing a novel, secured state channel. A characterization of blockchain peer-to-peer network itself in terms of size, geolocation and client software vulnerabilities, among others, is discussed in Section 5.4.14.

5.1.1 Link between T1.2 work and CONCORDIA pilots

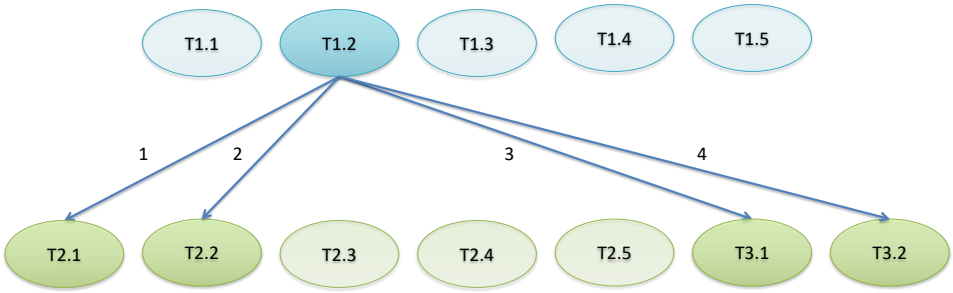

Figure 10: Relation between Task 1.2 and the CONCORDIA pilots

Figure 10 shows in which way the efforts undertaken in T1.2 relate to the pilots of the CONCORDIA project.

1. The extensive T1.2 work on data source acquisition and development provides a fundamental building block for CONCORDIA tasks in which actionable intelligence is required. The link between T1.2 and T3.2 is a strong example of this utility. T1.2's work on DDoS attacks has led to results that are being incorporated into the clearing house pilot. Concretely, in year 2, T1.2 members devised a novel methodology to classify infrastructure abused in reflection and amplification DDoS attacks. In coordination with T3.2, a prototype of this classifier is being used to enrich *fingerprints* within the clearing house architecture. Additional fingerprint metadata are planned for year 3;

2. The extensive investigation into DNS stability, security and resilience benefit virtually any application that depends on the correct functioning of the DNS. Transitively so, any CONCORDIA task in which such an application

is either being developed or a dependency stands to benefit. In addition, our investigation into DNS configuration issues (e.g., delegation, attack surface, weak points) provides actionable intelligence on threats. Efforts to operationalize sharing of DNS-related security incidents are underway. We will return to the importance of DNS-related research in our security roadmap considerations (see Section 5.2);

3. The studies on encrypted network traffic analysis and hardware-acceleration benefit CONCORDIA tasks in which threats can be learned from encrypted network traffic and traces (e.g., T2.1, T2.2 and T3.2);
4. The IP anycast related studies have high utility towards any task that stands to benefit from identifying services that are anycast, or assistance in enabling IP anycast deployments (e.g., T3.2);
5. The continued T1.2 efforts to bolster the blockchain technology and network are useful towards tasks that may levegage blockchain for sharing information on, e.g., threats.

5.2 Security Roadmap Considerations

Europe has an excellent track-record in the area of networks. Europe has played a major role in the standardization and development of mobile networks, with companies such as Alcatel-Lucent, Ericsson and Nokia. Technologies such as WiFi and Bluetooth were developed in Europe. Three of the largest Internet Exchanges are located in Europe (DE-CIX, AMS-IX, LINX), and connectivity for citizens and companies is world-class.

Europe is challenged, however, by the US and China (Huawei). If Europe loses control over its own networks, it runs the risk of becoming a digital colony of the US and China. Such development would not only have severe consequences for European companies (manufactures as well as operators), but ultimately our society and European values are at stake.

As Thierry Breton, the European Commissioner for the Internal Market, already said, the digital sovereignty of Europe rests on three inseparable pillars: computing power, control over our data and secure connectivity (i.e., networks). Whereas major European programs already exist for computing (processors, quantum) and data (GAIA-X), a major program for networking seems to be missing. In this section we will therefore identify some challenges to improve the security of European networks.

Probably Europe's biggest problem is that of fragmentation. Worldwide we witness a consolidation phase, where big companies take over smaller competitors. At this moment Europe has more than 50 mobile operators[35], of which only Deutsche

[35] https://en.wikipedia.org/wiki/List_of_mobile_network_operators_of_Europe

Telekom, Telefonica and Vodafone are within the top-ten[36]. The revenue of these three operators together is comparable to that of the biggest US operator (AT&T).

Because of this fragmentation, the security groups at most individual operators are relatively small and just able to follow the market. Real innovations often come from outside Europe, as is the case with DDoS protection services, DNS over HTTPs (DoH) and, more general, the collection of network data that may be relevant for security.

A long-term solution for these problems would be the consolidation of smaller EU companies into bigger, more powerful companies. Due to the federated nature of Europe such development would be politically extremely sensitive, and therefore not attainable on the short term. Fortunately, there are also a number of research and innovation actions that Europe could already take now to strengthen its digital sovereignty and to ensure the security and privacy of its citizens.

The key to all actions is break the monopoly of the big players by opening data and making infrastructures transparent. Next. we will identify challenges and possible actions to solve these challenges[37].

5.2.1 Open networking: The Responsible Internet

The problem of declining digital sovereignty is being addressed in several ways and in different areas of technology[38]. For example, Artificial Intelligence (AI) researchers have developed design guidelines to make the decisions of AI algorithms more transparent and explainable through what they call 'responsible AI'. Similarly, the European Commission is driving the development of a European federated cloud service called 'GAIA-X' that aims to improve Europe's data sovereignty. The European Commission recently also mapped out various policy instruments for areas such as 5G cellular access networks and the Internet of Things.

While these developments illustrate that digital sovereignty is a widely acknowledged and urgent problem, we observe the discussion largely overlooks the Internet infrastructure: the technical systems (e.g. routers, switches and DNS servers) that enable remote Internet devices to communicate with each other and that all of the other 'layers' (policy making, AI, data) depend upon. The exception is the debate around the alleged security weaknesses in 5G equipment. According to the European Commission, these pose a risk to the strategic autonomy of the European

[36] https://www.investopedia.com/articles/markets/030216/worlds-top-10-telecommunications-companies.asp

[37] Note: the term 5G security is sometimes used as umbrella to denote the various steps that Europe needs to take to make its networks secure. The problem with such term is that 5G is generally associated with mobile networks, leaving fibre and cable infrastructures aside. In addition, umbrella terms are generally not specific enough to identify the exact actions that need to be taken

[38] The text for this research challenge has been published in a web-blog before; the extended version can be found on: https://www.sidnlabs.nl/en/news-and-blogs/a-responsible-internet-increasing-trust-in-the-foundation-of-digital-societies

Union, but 5G networks only cover the cellular access part of the Internet infrastructure. The specific sovereignty problem in the Internet infrastructure is that users have no insight in, or control over how they depend on network operators and their systems, which ultimately poses a serious limitation for governments, institutions, companies, and individuals to decide how they can securely communicate. This is particularly relevant for critical service providers (e.g. power grids, transportation systems, mobile networks and manufacturing facilities), which have become increasingly dependent on computer networks. For example, such providers want to know if the Internet routes their traffic through networks with equipment that might have backdoors. At the same time, Internet users by design depend on third parties because the Internet is a massively distributed and global system of some 70.000 autonomous networks. For example, during a typical website visit, users unknowingly make use of the services of several DNS operators, transit providers, cloud services, and content distribution providers, all of which may reside in different geographical locations and jurisdictions

Actions To fill this gap in the digital sovereignty discussion, we propose the notion of a responsible Internet, a novel security-by-design extension of the Internet (or future networks) that offers users (e.g. providers of critical services or individuals) additional security-related options that give them a better grip on their dependencies on the Internet, thus increasing their trust in and their sovereignty over Internet communications. A responsible Internet accomplishes this by making its networks more transparent, accountable and controllable. This means users can ask a responsible Internet to provide high-level descriptions of the chains of network operators (e.g. ISPs, data centres and DNS operators) that potentially handle their data flows, for instance in terms of security and administrative properties, their interrelations and the management operations they carried out (transparency). A responsible Internet allows users to verify that these details are accurate (accountability) and to subsequently instruct the responsible infrastructure to handle their data flows in a specific way, for example by allowing them to only pass through network operators with certain verifiable security properties (controllability). The notion of a responsible Internet is inspired by responsible AI, a design paradigm that focuses on giving people more insight into how AI systems reach decisions and why.

5.2.2 Trustworthy DNS resolver infrastructures

The DNS system takes care of translating domain names (such as www.concordia-h2020.eu) into IP addresses (such as 139.91.90.171). Since DNS data provide a high-level overview of what network services exist and used, DNS data is crucial for security purposes. However, in the absence of proper privacy protection rules, DNS data can also be misused to monitor the behavior of individual users. Fortunately, Europe has strong rules to protect the privacy of its citizens

In the US such rules are lacking, and Internet providers are allowed to monitor the websites that their customers visit and sell that information to advertisement and

other companies. Since many customers don't like this, a number of US companies, most notably Google and Cloudflare, introduced the possibility to use DNS over HTTPs (DoH). By using DoH, Internet providers can no longer monitor the websites that their customers visit.

DoH is aggressively promoted by companies such as Google, and in the US various major browsers (i.e., Chrome and Firefox) use DoH by default. However, migration towards DoH introduces the following problems:

- US companies such as Google and Cloudflare are able to collect even more data of European citizens;
- For European Security Operation Centres (SOCs) and national intelligence services it becomes harder or even impossible to detect security breaches;
- One of the vital Internet services, i.e., the DNS, becomes under control of a small number of (US-based) companies. This introduces vendor lock-in and potential single points of failure.

Actions Although some aspects of DoH could potentially improve security, it is clear that changes are needed to solve the problems mentioned above. Research is therefore needed in the short term to address these challenges and make the necessary improvements.

5.2.3 DDoS protection Services

In a relatively short period the Internet has become one of the, or probably the most important infrastructure(s) that our society relies upon. If the Internet would fail, airports, harbors and shops would be closed, payment systems will fail, and working from home (in these times of COVID-19) becomes impossible. In the last decade we have witnessed an immense growth regarding the number as well as the strength of Distributed Denial of Service (DDoS) attacks on this vital infrastructure. Only 5 years ago most attacks were initiated by youngsters, spending a few Euros on a DDoS as a Service website (booter, stresser) to attack their favored bank. Fortunately, the mitigation of such attacks is relatively straightforward. Nowadays, however, we see ransomware attacks by criminals with strong technical skills on Internet and Service Providers. These new attacks are quite challenging and therefore have the potential to disrupt parts of our society for longer periods of time

To defend against DDoS attacks, many companies and organization have outsourced their protection to Akamai, Cloudflare and similar services. Although on average these DDoS protection services perform well, the fact that many of them are US-based create new problems. First, protection against application-layer attacks often require that these companies can decrypt all data, including sensitive data such as medical health records and online payments. In principle this gives Intelligence Services from outside the EU access to private information of EU citizens. This is not only undesirable, but might in some cases even be illegal. Second,

it creates a dependency of vital EU-services (such as healthcare and payments) on services from outside the EU. From the point of view of digital sovereignty, this is not what Europe should aim at.

Actions It is important to further develop open and European approaches towards DDoS protection. The DDoS clearinghouse, as being developed within the EU-CONCORDIA project, is a good first step. However, the focus of the DDoS clearinghouse is to share fingerprints of previous attacks, and not to protect against possible future attacks. Therefore, it is important to extend the Clearinghouse with protection capabilities. To cope with Terabit per second attacks, protection should be distributed over many locations, using technologies such as Anycast. In fact, a collaborative or federated protection architecture can be envisioned, in which similar services (for example banks or ISPs) share their DDoS protection capabilities to create a scalable DDoS protection service. More research on collaborative DDoS protection mechanisms is therefore needed now.

5.2.4 Monitoring and data collection infrastructure (data lakes)

The key to secure systems, services and infrastructures, is the availability of data. Examples of data relevant for (network) security include DNS data, BGP data, location data, log files, traffic traces (pcap and flows), open ports, etc. Data is not only needed to detect future threats, but also to understand trends. Data should therefore be stored for later analysis in so-called "data lakes."

Every day the Internet is scanned by many parties. For example, criminals scan to find potential ransomware victims, nation states scan to understand the state of the art, commercial organizations scan to share and sell data to interested customers. Examples of projects and organizations that scan the Internet include shodan.io, censys.io, RIPE Atlas and OpenINTEL. But also passive data is important for security; examples include BGP data from Hurricane Electric, RIPE's Route Information Services (RIS), traffic traces from CAIDA, and security incidents from the Shadowserver foundation.

Actions Europe should have the ability to collect, analyze and archive the data that it considers important to secure its citizens and society. Of course such activities should protect the privacy of its citizens by fulfilling the requirements of the GDPR, which means that a critical analysis is always needed to decide which data is collected, and which not. Such analysis needs to be transparent for the general audience.

From a research perspective, challenges include questions such as:

- how to perform scanning in a scalable and privacy sensitive way;
- how to quickly analyze huge data sets (big data analysis);
- how to correlate different and sometimes incompatible data sets (Machine Learning);

- how to condense and archive historical data, without losing precision;
- how to federate smaller data lakes to create bigger and therefore richer data lakes, without violating legislation or losing trust.

5.2.5 Network assurance & certification

The EU Cybersecurity Act introduces an EU-wide cybersecurity certification framework for ICT products, services and processes to ensure security and trust in ICT systems, including mobile networks, across development, deployment and operations. ENISA has a key role in setting up and maintaining the European cybersecurity certification schemes. For instance, ENISA is currently considering adopting the GSMA/3GPP NESAS/SCAS , certification scheme that has been jointly developed by GSMA and 3GPP for the certification of mobile networks equipment.

On the other hand, ICT technologies are developing at fast pace and rapidly introduced in ICT systems, that in turn are increasingly being developed and released and deployed following the Continuous Integration & Continuous Deployment (CI/CD). However, Security Assurance Frameworks (SAF) haven't evolved at the same pace as ICT systems.

- **Stasis** – SAF processes are defined for static targets with limited borders and features at a given point in time. Assurance for targets in development & operations is not sufficiently defined;
- **Slow and expensive** – SAF takes long time to conduct with human-based evaluation work by skilled experts from various security fields in addition to the target's domain of application;
- **Inertia** – upgrades or patches are either ignored or heavily delayed in domains with strict security SAF policies. Otherwise, vendors upgrade products but refer to outdated SAF proof;
- **Waterfall** – SAF follows conventional waterfall process whereas ICT systems are engineered increasingly by Continuous Integration Continuous Deployment (CI/CD) practices;
- **Blurred targets** – SAF is equipment/device-oriented for bundled software and hardware. But ICT softwareization and virtualization decouples software from infrastructure blurring the target's borders across software, infrastructure and service providers;
- **Technology (dis)trust** – there is a growing distrust on technology (origin) fearing backdoors in systems or components. It is not clear whether SAF can provide trustworthiness in this case;
- **Artificial Intelligence** – ICT systems are becoming AI-assisted. It is not clear how to evaluate AI unexplainable internals and its robustness against a new class of "intelligent" threats AI.

Actions In order to enable an agile and trusted EU digital market, where the latest technology can be leveraged in ICT systems that in turn can be trusted based on

evidence from agile security assurance frameworks, it is imperative to perform further research and foster innovate.

Short-term actions:

- Metrics: SAF should develop better quantitative metrics for measuring ICT trustworthiness;
- Explainability: SAF outcome is written for experts, but difficult to understand by stakeholders not in the security field. Explainable and comprehensive assurance is needed for legal purposes, business decisions and policy makers;
- Automation & formal proofs: SAF should leverage latest advances in AI for automation of the assurance and re-assurance process to reduce the human-factor that is subject to subjectivisms or prone to errors. Automation is also an enabler towards formal proofs of assurance.

Long-Term actions:

- Embedded: SAF should be agile and possible to embed in the ICT CI/CD lifecycle: development, deployment and operations. This would reduce the assessment and re-assessment burdens;
- AI: SAF shall include best practices end methodologies for evaluating the robustness of AI-based ICT systems that may contain bias or vulnerabilities against adversarial AI attacks;
- Softwarization & Virtualization: SAF should provide methodologies for assurance of virtualized and softwarized targets that are decoupled but still dependent on hardware and infrastructure.

5.3 Impact of COVID-19 on Research Activities

The practical execution of network-related research is largely 'terminal-based.' Specialized network infrastructures (e.g., flow exporters, programmable switches, and testbed setups) are, with few exceptions (for stringent security reasons), remotely accessible. In this respect, the pandemic has had limited impact on T1.2 members' abilities to execute the practical part of their network-centric research efforts.

From a viewpoint of coordination and management, meetings between T1.2 researchers from different institutions were already mostly on-line. In-person, physical events (CONCORDIA-organized and others) offered a welcome change of scenere at times. All of these events have gone virtual. This hasn't prevented T1.2 researchers from disseminating to other members (and tasks) their ongoing efforts. However, in-person gatherings and their social setting do provide a much better opportunity for ad-hoc discussions. Extended contact hours (in contrast to typically more short-lived and agendized virtual meetings) create a setting in which innovative ideas can flow. From this viewpoint, the COVID-19 outbreak has had a

degree of impact that is hard to quantify, but hasn't hampered T1.2-related efforts noticeably.

Many of the junior researchers in T1.2 are pursuing a doctorate degree. Some of them are not in their home country, removed from family, and have had limited to no opportunity to visit their office since Q1 2020, distancing them from their peers. Rather, they are confined to smaller (student) residences with a small social safety net. The impact of COVID-19 on the human side of is incalculable. Various T1.2 partners have indicated ways in which they try to weather this part as best possible, for example by organizing no business, social meetings on-line (coffee meetings, birthday celebrations, christmas events).

5.4 Summaries of T1.2 research efforts

The following sections contain more detailed write-ups of the research efforts that were introduced at the outset of this section.

5.4.1 Protection of the DNS infrastucture against DDoS Attacks

Contact: Raffaele Sommese (UT) Mattijs Jonker (UT)

Distributed Denial-of-Service attacks are one of the most disruptive attacks in today's Internet. These types of attacks are even more effective and more dangerous when they target or misuse core Internet infrastructure and services, such as the Domain Name System (DNS). The DNS is a fundamental pillar of the Internet's core infrastructure, and its role is crucial not only for the translation of human-readable names into IP addresses but also for supporting a myriad of widely used Internet applications, such as e-mail, VoIP, etc.

Within the frame of our research efforts, we try to understand what the weak point of the DNS infrastructure are, and which strategies are employed by network operators to increase the resilience of DNS. As part of these efforts we investigated the problem of delegation inconsistency between parent and child zone in DNS hierarchy. The DNS standard [93] states that the delegation records at both parent and child should be "consistent and remain so." However, we find evidence that this is not always the case. We dig into these inconsistencies and map the behavior of different popular resolver software with their default configuration to understand how they operate in these misconfigured scenarios. Moreover, we tested the behavior of the resolvers in the wild to obtain a clear view of the implication of these inconsistencies on the global Internet.

Our analysis reveals possible resilience problems, which could be leveraged by malicious actors. In particular, attackers could exploit the reduced resilience to lower the effort required to successfully perform Denial-of-Service attacks against DNS infrastructure. Moreover, we also find cases where these inconsistencies,

created by wrong management policies of DNS operators, arises potential risks of delegation hijacking.

The DNS is of vital importance to networked application. By assessing security properties of the DNS, we contribute to improved stability and resilience of the Internet.

5.4.2 TXTing 101: Finding Security Issues in the Long Tail of DNS TXT Records

Contact: Olivier van der Toorn (UT)

The DNS TXT resource record offers a lot of flexibility in terms of its contents, as it is a largely unstructured. Although it might be the ideal basis for storing any form of text-based information, it also poses a security threat, as TXT records can also be used for malicious and unintended practices. For example, TXT records can significantly add to the amplification potential of domain names in reflection and amplification DDoS attacks. Yet, TXT records are often overlooked in security research.

We present the first structured study of the uses of TXT records, with a specific focus on security implications. We were able to classify over 99.54% of all TXT records in a sizable dataset, finding security issues including accidentally published private keys and exploit delivery attempts. We also reported on our lessons learned during our large-scale, systematic analysis of TXT records.

Our accompanying paper [141] classifies a large portion of DNS TXT records to get a better grip on the security threats that may come from this core network infrastructure. This also benefits the applications that rely on the DNS, being developed within the context of CONCORDIA.

5.4.3 Preparing DNSSEC for Quantum Computing

Contact: Moritz Müller (UT and SIDN)

The DNS Security Extensions add integrity and authenticity to the DNS, the naming system of the Internet. Today, millions of domain names and users are protected with DNSSEC. DNSSEC achieves its goal by cryptographically signing content in the DNS and by recursive resolvers validating these signatures. In turn, this means only if the underlying cryptographic algorithms are secure, signatures can be trusted.

Today, operators, who want to sign their domain names with DNSSEC, have the choice between different algorithms. Each algorithm has different attributes in terms of security, key and signature size, and performance. One thing that all of these algorithms have in common is that a quantum computer could, potentially, break them in polynomial time. For this reason, the standardization organization

NIST is currently assessing algorithms that can neither be broken by current computers nor by quantum computers, so called quantum-safe algorithms.

In 2020, we studied if DNSSEC is ready for the challenge posed by quantum computers by trying to understand (i) how long it takes until new algorithms can be deployed in DNSSEC and (ii) if quantum-safe algorithms exist that are suitable for DNSSEC as well.

First, we analyzed and measured the complete DNSSEC algorithm life cycle, starting from the standardization in the IETF, to the replacement of insecure algorithms [99]. We showed that deploying a new algorithm at a significant number of domain names and gaining support at the majority of resolvers takes several years. Barriers include lack of support in software, lack of support in the registration channel, and reluctant domain name operators. Taking these barriers into account can accelerate the deployment of more secure and more efficient algorithms in DNSSEC.

Second, we studied if the quantum-safe algorithms, currently assessed by NIST, could be applied to DNSSEC as well [98]. Even if we can overcome the barriers highlighted in our first research, we still face additional challenges when replacing current algorithms with quantum-safe algorithms. All the assessed quantum-safe algorithms have either significantly larger keys, signatures or both. In our research, we defined requirements that quantum-safe algorithms should fulfill in order to be suitable for DNSSEC. Based on these requirements we picked three algorithms that we consider most suitable. Since these algorithms do not fulfill all requirements we proposed extensions and workarounds to the DNSSEC protocol.

Both studies combined lay the foundation for transitioning to quantum-safe algorithms in the future. In 2021, we plan to implement and test our proposals from the second research in realistic test beds. Thereby, we want to get a better understanding how these algorithms would perform in practice.

DNSSEC is being misused in DDoS reflection attacks frequently, due to its additional payload of keys and signatures. On the short term, transitioning to algorithms with smaller keys or signatures can reduce the attractiveness of the DNS for such attacks. On the long term, quantum safe algorithms pose a new challenge which need to be better understood.

5.4.4 Assessing anycast-based resilience of network services

Contact: Raffaele Sommese (UT)

Several studies [36, 94, 95] have shown that anycast is a common mechanism for network operators to increase the resilience of networked services. As the use of anycast to improve the resiliency of Internet services became common, and resiliency of critical communications infrastructure became a public policy issue, researchers have pursued methods for third-party inference of anycast deployment, i.e., identifying which addresses are anycast and from where.

Identifying address prefixes that are anycast enables a more accurate assessment of resilience properties of the Internet. As part of our efforts to understand the weak point of the DNS and which strategies network operators employ to increase DNS resilience, we devised a new measurement and inference technique to efficiently detect anycast prefixes. We call this methodology: MAnycast^2.

MAnycast^2 uses a distributed measurement platform of anycast vantage points as sources to probe potential anycast destinations, eliminating any sensitivity to latency dynamics that affect previous technique [135]. MAnycast^2 has proven to be efficient and scalable allowing us to complete, in under 3 hours, a full census of which publicly routable `/24` IPv4 prefixes are anycast.

Identifying resilience mechanisms and protecting network systems against DDoS attacks is a fundamental for network-centric security research. Our anycast inferrence methodology can be employed by other members of the CONCORDIA consortium in order to perform the same task on other core network infrastructures.

5.4.5 BGP Anycast Tuner: Intuitive Route Management for Anycast Services

Contact: Leandro Bertholdo (UT)

While Section 5.4.12 focuses on finding existing anycast deployments, the deployment of anycast is not easy per se. Big operators have their tools to monitor and configure anycast routing, but most of anycast networks are still configured manually.

We introduce a new approach to anycast management. Our solution is based on active measurements combined with traffic engineering. We propose the concept of a "BGP Cookbook" that allows operators to forecast the effects of routing policy changes over their services. We also introduce a web-based interface, called *"BGP Anycast Tuner"*, that allows operators to gain insight into their service's performance and provides easy management through automation. We evaluate our approach by implementing a prototype running in a testbed composed of 12 anycast sites covering 5 continents [17]. We demonstrate our tool in two different use cases: discovering and fixing a sub-optimal anycast routing issue, and shifting traffic between sites, which is useful during service disruptions as caused by DDoS attacks.

We propose to use previously mapped information of anycast catchment to help operators to address load balancing between anycast sites. We developed a tool, *"the BGP Anycast Tuner"*, that automates on the basis of measurements the configuration process of anycast routing. This tool allows a fine-grain control and administration of catchments, while provides an analysis and visualization of anycast management. *"BGP Anycast Tuner"* has an intuitive web-based interface, in

which operators can observe the predicted effects of routing policy changes, select an appropriate routing policy configuration, and deploy it automatically over tens of anycast sites.

The main contributions of our work are: (*i*) we establish a systematic approach to measure the distribution of clients of anycast services under different routing conditions, (*ii*) we present and release an open source tool, *"BGP Anycast Tuner"*, to manage anycast services using a routing *Cookbook* concept; and (*iii*) we evaluate our prototype in a real-world anycast testbed, compatible with a generic anycast network.

The results of the research efforts related to *"BGP Anycast Tuner"* can be integrated in the Threat Intelligence Center (T3.1), as our tool's output can be used to recommend mitigations actions to increase data availability and resilience when faced with DDoS attacks.

5.4.6 Looking Beyond the Horizon: Thoughts on Proactive Detection of Threats

Contact: Olivier van der Toorn (UT)

The Internet exposes us to cyberthreats attacking information, services, and the Internet infrastructure itself. Such attacks are typically detected in a reactive fashion. The downside of this approach is that alerts of an attack are issued as it is happening. We advocate that the security community could benefit by complementing traditional reactive solutions with a proactive threat detection approach, as this would enable us to provide early warnings by analyzing and detecting threat indicators in actively collected data. By describing three use cases from the DNS domain, we highlight the strengths and limitations of proactive threat detection and discuss how we could integrate those with existing solutions.

Our work was published in the Forum of Incident Response and Security Teams special issue of the Digital Threats: Research and Practice journal (DTRAP'20) [144]. Through DNS-focussed use-cases this research describes proactive threat detection from such a level that it directly applies to the Threat Intelligence efforts in CONCORDIA.

5.4.7 New Progress on Encrypted HTTP/2 Traffic Monitoring

Contact: Thibault Cholez (UL)

Encrypted HTTP/2 (h2) has been worldwide adopted since its official release in 2015. The major services over Internet use it to protect the user privacy against traffic interception. 77 websites of the top 100 websites use HTTP/2 nowadays, which makes HTTP/2 the future protocol for encrypted web traffic. However, under the guise of privacy, one can hide the illegitimate use of a service (for instance, according to an enterprise network security policy) and it is extremely important for a security practitioner to be able to monitor encrypted traffic. Therefore, the

analysis of encrypted network traffic is one of the main research objectives of CONCORDIA. While previous methods such as HTTPS proxies decrypt traffic on the fly to perform the analysis and by such expose users' private information, recent research works try to find a better security versus privacy trade-off by relying on machine learning algorithms.

In our prior work [21], we proposed a method using machine learning and able to identify user actions when browsing a website (for instance performing a keyword search). This solution relies on supervised machine learning using random forest and a proper set of features. It was a progress compared to the state of the art because of the high accuracy of the identification of HTTP/2 flows and the fact that only pre-defined and trained actions can be monitored, thus respecting user privacy. However, like many studies in the field, our study did not evaluate how generic is our solution to be directly applicable off-the-shelf (with no specific parameter customization per website) to any service and how long an acceptable accuracy can be maintained without retraining the model (to keep on with traffic changes in time).

The goal of our work this year was to answer these two questions. The main challenge here is the lack of availability of extensive datasets (in time and space) of HTTP2 traffic. Our work addresses these challenges by defining an experimental methodology applied on more than 3000 different websites and also over four months continuously. To summarize our main contributions are three-fold:

- Define a test-of-time and test-of-space methodology for HTTPS classification including the specification of a crawling campaign to collect relevant datasets. The collected dataset can be accessible on-demand (its size is too large to be publicly accessible on our servers). They consists of the pcap and HTML files but also screenshots of each accessed website;
- Evaluate our h2 classifier over around four months (test-of-time) and the impact of a regular re-training;
- Evaluate our h2 classifier on more than 3000 websites (test-of-space).

The results show that the method we used to classify h2 traffic is relevant and efficient for many different websites, assuming that the method has not been fitted to a particular one by design. However, the volatility of content in Internet as well as likely server software updates require regular re-learning, around every week in our case. This work has been published in the IEEE International Workshop on Information Forensics and Security (2020) [22].

5.4.8 Encrypted Traffic Analysis – TLS and QUIC Protocols

Contact: Pavel Čeleda (MUNI), Martin Holkovič (Flowmon)

Our joint research effort focused on flow-based encrypted traffic measurement and analysis of TLS (Transport Layer Security) and QUIC protocols. We continued work on the analysis of the encrypted traffic patterns of the TLS protocol and its changes introduced in version 1.3. The results are summarized in the article [84] and have been presented at IEEE/IFIP Network Operations and Management Symposium (NOMS) 2020. Extracted metadata from encrypted traffic (e.g., malicious domains and hostnames) are used to detect active cyber threats in network traffic. We have observed a massive increase in the volume and ratio of malicious encrypted traffic and COVID-19 related spear-phishing and malware campaigns in 2020. These targeted attacks focused on the Czech Republic's critical information infrastructure (e.g., successful ransomware attacks against medical facilities) or directly aimed at the Masaryk University's environment. We deployed and evaluated the JA3[39] fingerprinting to validate malware detection method based on fingerprinting of TLS client applications.

Next, we analyzed the QUIC protocol. QUIC is a general-purpose transport layer network protocol, and the HTTP mapping over QUIC is a new HTTP/3 standard. QUIC is a UDP-based, stream-multiplexing, and encrypted transport protocol that implements TCP-like properties at the application layer. QUIC combines HTTP/2 – multiplexing and flow control features, TLS – security, and TCP – reliability, congestion control and connection semantics. We developed a new Flowmon probe input plugin to process QUIC packets and to extract IPFIX information elements. The QUIC standard is still under development, and the production version of the input plugin supports the QUIC versions Q039 and Q040. The flow records are sent and stored on the IPFIX collector (ipfixcol2). We performed data analysis from the pilot deployment at Masaryk University, which used development version of the input plugin supporting arbitrary QUIC protocol versions. The majority of the QUIC traffic comes from the Internet and is generated by Google Services. The most frequently observed version of the QUIC protocol was Q050, the latest Google QUIC version, which has many similarities with IETF QUIC and is used by Google Services. The second one was the FAC2 version, which is an own implementation of QUIC protocol by Facebook.

Our work improves visibility into encrypted traffic and provides situational awareness of what is happening in the ever-evolving computer networks and new services. The CONCORDIA pilots may leverage the research results, e.g., T2.1 Telecom Sector: Threat Intelligence for the Telco Sector is considering sharing detected threats from encrypted traffic through the MISP platform (malware information sharing project).

[39] https://github.com/salesforce/ja3

5.4.9 Head (er) Hunter: Fast Intrusion Detection using Packet Metadata Signatures

Contact: Eva Papadogiannaki (FORTH), Sotiris Ioannidis (FORTH)

More than 75% of the Internet traffic is now encrypted, while this percentage is constantly increasing. The majority of communications are secured using common encryption protocols such as SSL/TLS and IPsec to ensure security and protect the privacy of Internet users. Yet, encryption can be exploited to hide malicious activities. Traditionally, network traffic inspection is based on techniques like deep packet inspection (DPI). Common applications for DPI include but are not limited to firewalls, intrusion detection and prevention systems, L7 filtering and packet forwarding. The core functionality of such DPI implementations is based on pattern matching that enables searching for specific strings or regular expressions inside the packet contents. With the widespread adoption of network encryption though, DPI tools that rely on packet payload content are becoming less effective, demanding the development of more sophisticated techniques in order to adapt to current network encryption trends.

Our efforts have been focused on the implementation of HeaderHunter, a fast signature-based intrusion detection system even in encrypted network traffic. We generate signatures using only network packet metadata extracted from packet headers. Also, to cope with the ever increasing network speeds, we accelerate the inner computations of our proposed system using off-the-shelf GPUs [109]. The next figure presents the performance achieved by our pattern matching engine. More specifically, we show the throughput sustained by the discrete GTX 980 GPU executing the pattern matching engine of our intrusion detection solution.

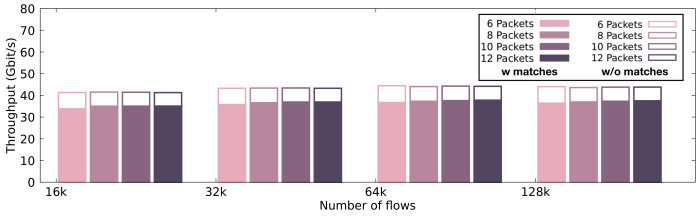

Figure 11: Performance achieved by our pattern matching engine

In Fig. 11, the color-filled bars indicate the performance achieved by the pattern matching engine when the selection of (i) signatures and (ii) input results to a computationally loaded condition. In the figure, we present the worst-case scenario, where we have full contamination of the traffic. White-filled bars with borders indicate the performance achieved in a computationally relaxed condition (i.e., less than 10% infected traffic), which is the most realistic scenario. We present the

throughput using different packet batch sizes. This works applies to the network-centric security in CONCORDIA.

5.4.10 Threat Intelligence at IoT Edge (both encrypted/unencrypted traffic)

Contact: Han Wang (RISE), Shahid Raza (RISE)

With the introduction of new privacy laws, such as the General Data Protection Regulation (GDPR) in the EU, the amount of data shared should always be minimized. Federated Learning (FL) is regarded as a promising solution to this issue because it provides strong privacy protection for the participating entities as data has not to shared directly. Privacy-preserving: We apply FL at the edge of IoT networks, where the actual collection and/or actuation of data takes place. This is becoming a major trend. Its advantages include not only privacy but also efficient resource utilization and reduced latency. However, as IoT devices usually are resource-constrained compared to cloud-based systems it opens up research problems on how to, in a resource efficient way, implement the FL components in the edge of the IoT network.

This year, we have applied federated learning to the problem of IoT device fingerprinting and identification. The results are documented in a paper, which is under review in USENIX Security Symposium 2021. We analyze IoT network traffic data to develop a fingerprinting mechanism to generate lightweight fingerprints used in the device identification scheme. The proposed method can be applied with both encrypted/unencrypted traffic. The results in this paper has shown the effectiveness of federated learning. Despite the advantages brought by FL, there still exist challenges. For example, the heterogeneity of IoT devices implies that collected datasets usually are class-wise imbalanced and non-IID (Independent and Identically Distributed), which generally causes the FL model to be biased in favor of well represented classes, and at the same time, the performance is degraded. We aim to solve this problem in the upcoming research work, and we will also look into how to develop mechanism against adversarial attacks such as poisoning attacks in federated learning.

5.4.11 A Measurement-Based Approach to DDoS Mitigation

Contact: Mattijs Jonker (UT)

There are many challenges when it comes to DDoS mitigation. Knowing what it is exactly that we are defending against constitutes a challenge. Understanding the uptake and operation of (commercial) mitigation solutions is another. Moreover, acquiring and developing (raw) data sources to methodologically study the DDoS problem constitutes a challenge in itself. We contributed significantly towards overcoming these challenges by taking a measurement-based approach. We used large-scale passive and active measurements from diverse vantage points all over the world, to gather a variety of independent data types [75].

By successfully fusing diverse data we: *(1)* unveil eye-opening statistics about global attack activity; *(2)* gain insights into the Internet-wide adoption of mitigation solutions as well as operational practices of users; and *(3)* lay bare and investigate the undesirable side effect of mistakes in deployment and operation. We also made some of the resulting data available to the research community.

We present a large-scale characterization of attacks and reveal the massive scale of the DDoS problem. We also advance our understanding of the adoption and operation of mitigation solutions and reveal global trends in adoption as well as operational practices. Finally, our work underpins that mistakes are made in deployment and operation, which arguably leave some operators and users with a false sense of security. Our work also corroborates the notion that attackers can seize on such mistakes as an opportunity to bypass defenses. Our research efforts enabled us to inform policy makers and regulators dealing with societal questions, in addition to the research community and network operators. This gives meaning to the work beyond its scientific contributions.

5.4.12 Generation of Encrypted Network Traffic Datasets

Contact: Martin Drasar (MUNI)

In the last year, our efforts in generating encrypted traffic datasets were focused on the ability to create semi-automated datasets from preexisting network traffic traces. These efforts culminated in creating and open-sourcing the Trace-Share utility [3]. In addition to Trace-Share, our research was focused on cybersecurity interaction simulation environments. This year the simulation environment gained research prominence as a vehicle for producing datasets of non-trivial attacker behavior, tapping into the resources provided by Trace-Share and other platforms.

The core thesis of our research into cybersecurity simulation platform is that historically, cybersecurity research on adversary behavior was reactive rather then proactive. However, the advent of advanced persistent threats (APTs) brought into forefront that malware is sophisticated, custom-tailored and stealthy. As such, it rarely provides timely and usable datasets for post-mortem analysis of attack tactics and methods. We assert that cyber defense requires a switch from the reactive approach to proactive, wider adoption of simulation and emulation techniques that include active exploration of novel adversarial methods and strategies. Only then we will acquire datasets that are current and comprehensive.

While the simulation and emulation techniques have been used for decades, most have a narrow scope, focusing on modelling the impact of attackers' actions and their mitigation but not on the adversary actions and strategies. Others have a significant overhead to enable massive simulation that require training of actors based on machine-learning. To this end, we aim to fill the gap by simulating cyber adversary actions with a focus on their attack strategies. We are developing a dis-

crete event cybersecurity simulation framework, CYST, which fulfils the following goals:

- lightweight simulation of multi-agent cybersecurity scenarios;
- streamlined integration with machine learning toolkits;
- integration of different attack, defense, and observation models;
- hybrid-stochastic simulation of interaction between attackers and defenders for in-depth analysis of attack strategies;
- rapid prototyping of attack and defense strategies;
- smooth transition of simulated actors into emulated and real-world settings;
- modelling of environments, which can be emulated in virtual environments using already provided data;
- integration of simulation and emulation to remove the need to re-implement existing cyberdefense mechanisms.

A preliminary version of CYST was already presented [44] and made available as a snapshot of its functionality [1].

Currently, CYST is being extended with three capabilities, which will play a crucial role for the ability to generate quality datasets, both encrypted and unencrypted:

- **Automated generation of realistic cyberscenarios**. A necessary feature for any robust simulation and training of ML-based agents, as well as for exploring different scenarios. We already have an implementation based on describing scenarios as a set of constraints and treating it as a satisfiability problem;
- **Implementation of immediate and delayed observation models**. Converting request-response level messages within simulator to observations serves as a basis for defenders' decisions and for creation of network traffic datasets. These observations will support different levels of abstraction (event, flow, packets) and different operation modes (host-based, network in-line, network behind TAP);
- **Implementation of defender models**. There are different approaches to defense, which need to be modeled and implemented to provide realistic capabilities. These include passive vs. active defense, localized vs. infrastructure-wide defense, usage of honeypots, etc. These different capabilities are a key to creating response baselines within datasets.

In the coming year, we expect to focus on making these capabilities accessible to general public by means of releasing another version of CYST framework. We also

expect to employ the aforementioned capabilities in collaboration with the TraceShare utility to produce datasets of non-trivial adversarial behavior. These datasets will be evaluated and also made available.

5.4.13 Tackling the scalability issues of blockchain networks by the incorporation of state channels based on distributed and decentralized web

Contact: Blaž Podgorelec (UM), Muhamed Turkanović (UM)

Blockchain technology attracted a lot of interest from the general public and various enterprises that want to exploit it to support their existing business processes or develop business models. However, some drawbacks are identified when considering the most stable and popular blockchain platforms. One of them, and probably the main, is certainly scalability. Many techniques exist on how to tackle the scalability issues. One is to transfer a state outside the blockchain network and introduce an off-chain state channel, where the state can be manipulated without the blockchain network's built-in limitations. Furthermore, the state channel allows that, at some point, the off-chain state is transferred and merged back to the blockchain network, i.e., on-chain. Still, the state channel solutions face the challenge of preventing users' violations, specifically transferring the invalid off-chain state to the blockchain network. Moreover, the challenge is also how to transfer the off-chain state into the blockchain network in an ad-hoc manner. To solve the problems mentioned above, we propose a novel state channel solution in the form of a state channel as a service [111], which, although off-chain environment, still incorporates a secure distributed and decentralized network (IPFS). The proposed solution solves the challenge of transparency and traceability while giving users the assurance and ability that in an ad-hoc manner, only the last off-chain valid state is transferred back on the blockchain network.

Our work's contribution improves the network security between the two different environments, i.e., on-chain (blockchain network) and off-chain (state channel), while facilitating the scalability of solutions built on top of blockchain technology. Moreover, the security and the usability for users that are part of the state channel are also increased since there is no need to monitor potential malicious activities from other parties, and the ability for an ad-hoc transfer of state on-chain is possible.

5.4.14 A Comprehensive Study of the Bitcoin P2P Network

Contact: Thibault Cholez (UL)

During the last decade, cryptocurrencies have gained in popularity. Even if their total market capitalization is prone to much variation, they are currently valued to more than $480 billion. With more than 18 million of circulating tokens valued to approximately $320 billion, Bitcoin is the most acknowledged cryptocurrency. The Bitcoin blockchain is managed by an underlying peer-to-peer network. This

network is responsible for the propagation of transactions carried out by users via the blocks (which contain the validated transactions), and to ensure consensus between the different nodes. The quality and safety of this network are therefore particularly essential as it became an essential assets that can be the target of cyber attacks.

Due to the popularity and success of Bitcoin, its protocols, network and security have been widely studied. The Bitcoin community has used these research to develop fixes and countermeasures to these attacks. But, it lacks thorough studies of the network as a whole, and the ensuing implications. Last year, we performed a state of the art on the subject that was reported in D1.1. To the best of our knowledge, even if some authors indicate their methodology, either their dataset are not publicly available or the source of the software used (crawlers, scripts to build the statistics, algorithms, ...) are not given, preventing reproducible results. None also proved that the crawler used converges and is sound. Beyond methodology questions, some important criteria are yet unknown like nodes popularity in contact lists or the state of the network regarding software vulnerabilities. Finally, those studies were performed before the introduction of limiting countermeasures, and their impact on our current measurement capacity, specially regarding link inference, must be evaluated.

Our research work this year consisted in the creation of a proper dataset composed of snapshots of the Bitcoin network and conducting a comprehensive analysis of it. During one month, we performed regular crawls on the nodes composing the network and gathered information about them. Our dataset and analysis are made by our open source crawler and scripts and thus fully reproducible. We highlight some metrics that characterize the network. Among these metrics, we analyze a few classical ones like the size of the network, the geographical localization of peers and the churn to get fresh results, but also new ones, in particular the popularity of peers, and the inventory of software vulnerabilities that affect clients' versions and their distribution among the deployed nodes. We also show that link inference is still possible despite the added countermeasures.

Our results show that the size of the P2P Bitcoin network is very stable. The peers composing the network are well balanced throughout the world and show little churn. But the network also exhibit more concerning properties like the fact that a significant part of the network tends to update the client version slowly, and the unbalanced popularity of peers, even among the reachable ones. Our latest experiments show that link inference is still possible. This work has been submitted at the IFIP/IEEE International Symposium on Integrated Network Management. Next, we will discover and maintain a view of the network topology and analyze its dynamic graph properties to identify possible weaknesses.

6 Software/System-Centric Security (T1.3)

Task 1.3 (T1.3) of the CONCORDIA project is concerned with software/system-centric security. Specifically, the main research topics addressed by T1.3 are:

- malware analysis;
- security by design: adaptive software and operating systems;
- detecting service dependencies; and
- system security validation and zero-days.

6.1 Overview

A summary of T1.3 contributions are given in the following paragraphs focusing on year 2.

The CONCORDIA approach to *security by design: adaptive software and OSs* is to provide a run-time and adaptive *on-demand* security. Studies are focusing on (a) identifying threats as they occur, (b) analyzing the propagation of the attacks, and (c) providing run-time mitigation security ?enhancers? on-demand. ULANC partner works on preventing attacks, in particular in cloud systems, by identifying the attack surface perimeters . For this, they corroborate various vulnerability data using essentially two approaches : ontologies and Petri-Nets. RISE partner works on software security in IoT devices. For this, they conducted research on hardware-assisted Trusted Execution Environments (TEEs) that are implemented in IOT devices. In particular they focus on TrustZone-M.

On *Malware analysis*, ISI partner develop malware detection heuristics based on machine learning. For this, they use static analysis to extract various features and in particular the graph of API calls. UL partner works on malware analysis based on dynamic analysis. BD worked on improving the file-less and malicious documents detection based on the feedback from exploiting the techniques that were proposed last year. UL Partner with the start-up CYD is continuing their collaboration on developing new approach top detect malware. The tool named Gorille is now accessible on demand at . Finally, we provide a BLOG input on malware analysis .

A goal of *Detecting Service dependencies* is to analyze performances of services. ULANC partner works on performance testing in order to identify performance bugs. A key challenge of performance testing is that injecting performance bugs must not alter the functional behavior of the original program. A framework called SlowCoach to perform performance mutation testing (PMT) has been developed.

In *System security validation and zero-days*, an important issue in software maintenance and long-term support is to be able to know whether or not a (new) common vulnerability and exposure (CVE) is present on deployed commercial, off-the-shelf

software components (COTS). In order to achieve this important issue, a step is to identify toolchain provenance. UL partner combines formal methods and machine learning to determine the compiling chain used to generate a given bare binary code.

6.2 Link between T1.3 work and CONCORDIA pilots

The research performed within Task 1.3 is linked to the CONCORDIA pilots as follows.

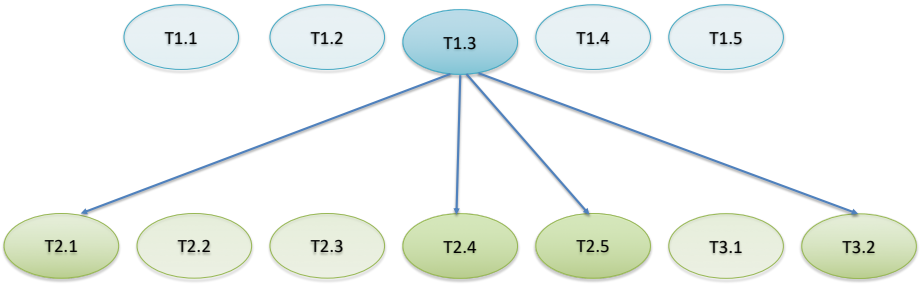

Figure 12: Links between task T1.3 and the CONCORDIA pilots

BD contributed to the telecom pilots with the collaboration with Telecom Italia. BD contribution was on the data validation and augmentation with detection information. The collaboration is still active and it will continue to be visible through the MISP information enrichment. Discussions have started to also provide (malware) samples. IC are using its works in the UAS pilot and are working in conjunction with Airbus on a scenario to analyse resilience in the context of heterogeneous fleets of aerial vehicles (for details see D2.2).

The other links are still under discussions and we expect it will progress soon.

6.3 Software/System-Centric Security Roadmap Considerations

In following section, a roadmap is developped. The two first challenges are the current ones. The other challenges are future research foreseen in ordre to improve the security of systems includes research on Quantum Technologies and Artificial Intelligence.

6.3.1 Malware detection and analysis.

Ransomware, and more generally malware encompassing a lot of other threats like spyware and botnets weaken our digital systems. The surface of attacks of malware are broader and broader, it includes all IT infrastructures: computes, smartphones and tablets, IOT devices, cars, and industrial infrastructures. They are aimed at

the ordinary citizen as well as companies and administrations, even hospitals. The design of these malicious codes is increasingly complex. That is why even old malware strains can be undetected, like recent Emotet attacks. The consequences are financially huge and can can also lead to a malfunction of our critical infrastructures.

Actions: In this arm race, it is necessary to develop new malware defense concepts. An holistic approach that takes into account a broad set of information, is necessary. That said, there is also a room of improvement to devise new cutting-edge anti-virus products by combining machine learning and formal methods along with system events augmentation. *Lastly, it is crucial to have access to shared platform of malware collection and their related information.*

6.3.2 Service Dependency Roadmap

The complexity and a plethora of services involved in distributed systems such as the Cloud entails significant and often manual work to understand the interconnection and the behavior of the services in the system. This hinders the profiling of threats and their propagation in the system. We plan to automate this process by using the capabilities of model checking that would essentially enable profiling and analyzing the potential paths that could be taken by a threat to propagate in the system.

Actions: The midterm goal for the service dependency task is to develop techniques to perform automated multi-level threat detection in large-scale data center/Cloud systems. This inherently enables the Cloud providers to assess the potential propagation paths of the threat and consequently, prioritize the services accordingly.

6.3.3 Explainable Security deep analysis

Nowadays, Machine Learning (ML) and Arfificial Intelligence (AI) approaches are more and more prominent as methods to analyze, classify, and then take actions. This is quite well-known in systems like face recognition, but there are other applications like network traffic analysis or malware detection. In each case, it is important to be able to explain an analysis performed by AI systems and give reasons justifying actions taken. Thus in Forensic, proofs or attribution of an attack is a key issue, and so an analyze should be returned enough explanations. Another field is the one of embedded systems. Decision systems in a car should be able to provide a reason of a decision.

Actions: In the domain of cyber-security, it is worth to develop Explainable Security deep analysis. This domain is already an important subject in AI, so we should have a closer loop in this direction.

6.3.4 Quantum Technology

Quantum Technology (Q-tech). Q-tech is receiving high attention in research, industry and governmental agencies. It is therefore important to outline an informed strategy based on a good understanding of the current status of the Q-tech and prioritize the right topics.

Based on existing research in Q-tech related initiatives we can summarize the current status as follows:

- Quantum Computers - building a quantum computer is highly expensive and difficult. Its application is not general yet, i.e., they can efficiently solve few specific problems.
- Quantum attacks on crypto ? A recent report by experts from academia and industry judged that the construction during this decade of a quantum computer capable of breaking currently used public-key crypto would be highly unexpected. Symmetric crypto is quantum-safe, e.g., SIM card authentication. The business case for quantum adversaries is thus questionable. However, quite a lot of research and development is focused on post-quantum cryptography (sometimes referred as quantum-proof, quantum-safe or quantum-resistant).
- Quantum crypto - Evaluating and standardizing new crypto-systems necessarily takes time. The industrial benefits of quantum crypto are not directly applicable to all industries. Each industry sector needs to assess its suitability and feasibility.
- Quantum key distribution (QKD) ? QKD is suitable in quantum communications and research shall remain in this quantum domain. QKD is useless otherwise as a replacement of currently established key distribution protocols used for authentication, signatures or integrity.
- Governmental intelligence agencies ? Based on authoritative sources, they are not in a hurry replacing commercially used public-key encryption.
- Quantum simulators ? while useful in some domains, quantum simulation environments for cybersecurity purposes are questionable and no meaningful use case has been identified.

Actions Based on the current state of the art and estimations about the expected progress the following research is needed:

- Open post-quantum crypto: Research in post-quantum crypto (aka quantum-safe) is of high-importance including wide and active participation in relevant standardization bodies such as IETF, NIST, 3GPP in order to ensure many-eyes expert reviews in an open transparent process. We need to avoid lock-in proprietary schemes taking over the market.

- Resilience: For industries relying on public-key cryptography (PKC), prepare risk-based recommendations on: (i) develop post-quantum systems based on authoritative upcoming NIST standards; (ii) prepare timed transition processes based on the progress of the authoritative research community; (iii) prepare replacement, contingency and containment strategies. For industries, this includes inventories of PKC-based protocols used (TLS, IPsec, S/MIME, SSH) and its base deployment in devices, appliances, networks and services.

6.3.5 Adversarial Artificial Intelligence attacks and countermeasures

A very important aspect to be considered in AI usage for security purposes is the intrinsic vulnerability of AI data, algorithms and models to adversarial AI attacks. This new attack surface can be considered hard to mitigate. AI adversarial attack cannot be fixed since they rely on the learning nature and unavoidable use of data of an AI algorithm. AI technologies can be used as weapons for performing cybersecurity attacks by generating malicious traffic, malicious code as well as automating the hacking process. This weaponization of AI can be very potent since it is adaptable to the countermeasures provided by defenders. In parallel to this type of attacks, data poisoning and model poisoning can also be performed in to order to attack an existing AI infrastructure. This adversarial attacks on legit AI systems aim to render such system blind to specific type of inputs or reduce the AI systems? accuracy as a whole. The current threat landscape is very broad and has been identified as critical for the secure use of AI in European security/privacy sensitive domains (Law Enforcement, Health, Critical infrastructure domains etc.). Also, is should be mentioned that there exists no well-structured detection framework that can assess vulnerabilities of AI systems against adversarial AI attacks or weaponized AIs. Given the growing usage of AI solutions, the need of such an assessment mechanism becomes great.

Actions Acknowledging the potency of the above-mentioned attacks, agencies, organizations as well as industries across Europe should establish a ?security net? for detection, response and mitigation. The Goal should be to create the means in order to: 1) reduce the risk of attacks on AI systems, and 2) mitigate the impact of successful attacks. AI adversarial attack protection (security net) can be structured in three layers, planning, implementation and mitigation.

- Planning: At the design phase of an AI solution, including evaluation of possible training datasets as well as choice of AI classifier and modeling algorithms, an AI risk assessment process could be formalized to perform ”AI Suitability Tests? that assess the risks of current and future application of AI datasets and algorithms. An acceptable level of AI use within a given application could be provided as an outcome. These tests should weigh the application?s vulnerability to attack, the consequence of an attack, and the

availability of alternative AI-based methods. Apart from the above, the AI risk assessment ca also perform a formal validation of data collection practices and suggest mechanisms for protecting data and restricting data sharing to trusted entities only. Finally, in the planning layer, best practices should be extracted so as to manage the entire lifecycle of AI systems in the face of AI attacks. These practices apart from technical aspects they will include strategic, operational as well as legal/ethical aspects of AI deployment.

- Implementation: During this layer, the best practices should be further consolidated into adopted IT related reforms on ATI solutions so as to make AI attacks more difficult to execute. The process relies heavily on setting up security/cybersecurity mechanisms that will protect the assets which are used to craft AI attacks, such as datasets and models e.g by improving the cybersecurity of the systems on which these assets are stored. This includes installing cyber defense mechanisms that support the CIA triad and detect cyberattacks (intrusion detection, anomaly detection etc) using hardware and software means.
- Mitigation: Mitigating AI attacks is not an easy task since such attacks are advanced and have very recently appeared in the security domain. Existing research proposals should be extended to mature solutions. Detection and Mitigation techniques could rely on decreasing the success rates of backdoor (harder to identify and track) attacks also known as poisoning attacks (e.g ?pruning method?) but also techniques that introduce defense mechanisms (for detecting AI attacks) like Adversarial Training, Defensive Distillation, Generative Models and Regularization of datasets. The goal of the mitigation layer should be to:
  - harden AI models to be resistant to fault data injection and poisoning attacks (during design)
  - infuse the AI models with detection mechanisms so that they can classify (apart from valid data) also malicious data (during AI operation).
  - record the cybersecurity incident related to the detected attacks and report it to the cybersecurity community.

6.4 Impact of COVID-19 on Research Activities

The impact of COVID-19 on research work is relatively minor. That said, the consequences are rather on the constitution of a research community T1.3. The fact that all planned physical meetings had to be canceled was a hard blow. Indeed, with the change of task leader last year, we are still trying to get to know each other better in order to work together. Another impact of COVID-19 is that we were not really able to exchange on emerging topics that we wanted to push such as IOT malware. Hence, after the first main lockdown (spring 2020), we decided to start

monthly online meetings. We are now a little better organized and for example a collaboration is starting between BD and UL partners on malware analysis.

6.5 T1.3 Highlights

Four partners, BD, CYD, ISI and UL, are addressing the challenge in the fight against malware. This research is the beginning of a collaboration between the two companies, BD and CYD, on one hand and two academic players ISI and UL. The development of Gorille Cloud made jointly by CYD and UL is a promising solution to tackle cybercrimes. We are also working on machine learning methods for malware analysis which can be considered as emerging solutions, but which have yet to show their capacity in real situations. Another highlight is the works of ULANC and RISE partners addressing the challenging of investigating vulnerabilities in complex IT environment including Cloud infrastructure and IOT devices.

6.6 Summaries of T1.3 research efforts

The sections below contain an account of the diverse research activities of task T1.3.

6.6.1 Security by design: adaptive software and OSs

Contact: Salman Manzoor (ULANC) Anum Khurshid (RISE)

The enumeration of attack surface is a challenging task given the plethora of services and interfaces involved in ecosystems such as the Cloud. Therefore, the current direction of research at Lancaster University (ULANC) focuses on the usage of vulnerability data to establish a context of the attack by incorporating diverse features of the vulnerability. This context is used to reason on the further attacks targeting the system. Moreover, we are developing a new multi-level approach to identify the threat propagation across the services involved in the Cloud operations.

Ontologies for Vulnerability Terrain Mapping and Attack Reasoning

Contact: Salman Manzoor (ULANC)

The characteristics of cyber-attacks (e.g., nefarious IP addresses, malware hashes) are relatively easy to detect and blacklist. However, to perform a comprehensive analysis, establishing a ”context” is required to facilitate system defenders in understanding the manifestation of the vulnerability across the system. Furthermore, the context assists in exploring common attack characteristics among the vulnerabilities. For instance, revealing attack mechanisms prevents attackers from exploiting the same attack mechanism across vulnerabilities. We argue that an ontology-based approach may be followed to help in building context to explore ”vulnerability terrain” based on the contextual similarity among the vulnerabilities. To build the context, we utilize and extract data from the National Vulnerability Database

(NVD) to map the data to the respective ontology class(es) which can be partially automated for usability. We perform reasoning on the ontology to identify different characteristics of the vulnerability terrain. For instance, our work shows that recurring actions operated by an attacker in a Cloud environment are associated with manipulating the combination of legitimate Virtual Machine (VM) actions that lead to an adverse impact on the Cloud functionality.

VulPro: Profiling and Exploring Threat Propagation in Cloud Services

Contact: Salman Manzoor (ULANC)

Traditional Threat Analysis (TA) approaches have been applied to identify potential threats in distributed systems such as the Cloud. However, TA approaches for the Cloud typically focus on a single service or on a particular technology and therefore, TA based on a single service is incomplete as they fail reveal the threat propagation across the boundary of the service. Moreover, knowledge of threat propagation patterns and unsuccessful execution of security defense protocols is important for more accurate and detailed threat analysis to enable (a) exploring threat progression in the system and to identify the ”common” paths associated with the threat propagation, (b) analyze the shortest path to attacker's objectives, and (c) analyze the impact of multiple threats across different services in the Cloud. We present a new approach – called VulPro – that aims to capture threats across different services involved in the Cloud by leveraging Petri nets to design a Multi-level modelling of the system and the vulnerabilities. The obtained model sets the basis to comprehensively generate a computational tree to enumerate the progression of the threat and identify common paths leading to different attacks. We validate the model using data published in national vulnerability database.

Trust on IOT infrastructures

Contact: Anum Khurshid (RISE)

Our Efforts in task T1.3 began last year with exploring issues that hinder software security in IoT devices, identifying and developing technologies that guarantee trust on IoT infrastructure including the underlying architecture, system components and data. Our approach to attain this assurance and security of software components is through isolation of critical components and data using hardware-assisted Trusted Execution Environments (TEEs). We focus our research on TrustZone-M, which is a newly introduced TEE for resource-constrained IoT devices and is still undergoing development and challenges. In the previous year, we proposed a framework for establishing a secure communication channel between the non-secure and secure domains of TrustZone-M. This work has resulted into a research paper: ShieLD: ”Shielding Cross-zone Communication within Limited-resourced IoT Devices running Vulnerable Software Stack” and is under submission in IEEE Transactions on Dependable and Secure Computing which is a reputable security journal.

We continue our work with TrustZone-M with identification of another research problem. In the current setup of TEEs, software/system components in the secure side have access to the entire system resources. This calls for device users to trust the secure software counterparts of the applications originating from different vendors. The device manufacturer has to ensure that the hardware components/peripherals and the accompanying software do not perform unnecessary access or abuse of resources that it can access. Installing rootkits and illegal file sharing using compromised secure zones are some potential ways TEE-enabled systems can be exploited. The fundamental problem is that the IoT market is diverse, unregulated and still expanding. Hence, we identify the need for a solution that ensures that peripherals and their software incorporated into secure areas on a TEE-enabled device do not misbehave and deviate from their intended functionality. This is an ongoing work, where we are designing a solution to monitor secure software and the way they access secure on-device peripherals.

Evaluating the resilience of systems

Contact: Emil Lupu(IC)

We have been in particular investigating methods for evaluating the resilience of systems. By resilience we mean the ability to preserve the operation of the system even when an attack is in progress and some parts of the system may have been compromised. This requires to model both the progression of the attack in the system and the impact on the operation of the system that a partial system compromise can have. We typically use attack graphs to represent the progression of multi-step attacks, where the attack moves laterally from one system to another as well as within a single system (privilege escalation). The model used to represent the impact of an attack on the system operation may vary depending on the system and the type of analysis considered. Dependency models are a straightforward way to represent and reason about impact on system integrity. Performance models such as queuing networks can be used when it is necessary to reason about system workload and offer a finder grained analysis in some circumstances.

6.6.2 Malware Analysis

Contact: Dimitrios SERPANOS (ISI) Jean-Yves Marion (UL) Régis Lhoste (CYD)

Weaknesses of machine learning systems for malware detection

Contact: Emil Lupu (IC)

We are also currently collaborating with CODE in the investigation of the weaknesses of machine learning systems for malware detection to attacks at run-time. It has been shown that machine learning algorithms are very vulnerable to adversarial examples, where attackers can produce intentional errors at run-time by introducing small perturbations to original data points. In our context of application,

malware developers can introduce additional features capable to evade detection by machine learning algorithms. Our research focuses on understanding the systemic vulnerabilities of machine learning algorithms in this application domain to Universal Adversarial Perturbations (UAPs), where the same modification applied to a large set of malware examples is capable of evading detection for a large fraction of them. This poses a serious risk, as the effort required by attackers to evade the system reduces and enables the possibility of creating tools for modifying the malware, preserving the malicious functionality, to evade detection by machine learning components in antivirus systems. We are investigating the impact of UAPs both in the feature and the problem space, exploring both Windows and Android malware, and developing new defensive techniques to mitigate this threat. .

Malware Classification based on Abstract API Call Graph Method

Contact: Dimitrios SERPANOS (ISI) V. Tsouvalas (ISI), V. Pikoulis (ISI)

We developed a graph-based solution to the malware detection problem, which implements resource extraction from executable samples and applies machine learning algorithms to those resources in order to classify the executable as malicious or benign. Given an unknown Windows executable sample, we first extract the calls that the sample makes to the Windows Application Programming Interface (API) and we construct an API Call Graph, which is converted to an Abstract API Call Graph. Subsequently, using a Random Walk Graph Kernel, we quantify the similarity between the graph of the unknown sample and the corresponding graphs originating from a labeled dataset of known samples (benign and malicious Windows executables), in order to carry out the binary classification using Support Vector Machines. Following this method, we achieve accuracy levels up to 98.25%, using a substantially smaller dataset than the one proposed by similar research efforts, and being considerably more efficient in time and computational power. The work is currently prepared for submission.

Malware Classification with Machine Learning Techniques

Contact: Dimitrios SERPANOS (ISI) G. Xenos (ISI)

Malicious software poses an increasingly important threat to the security of computer systems. A modern approach for its detection and classification is the use of machine learning models. In this work, we developed an automated detection model that uses data derived from static analysis of malicious software. Furthermore, we explored various techniques for the construction of such models, in order to evaluate their capabilities and to make useful conclusions about malware detection. Specifically, the model is based on EMBER, a labeled benchmark dataset for training machine learning models, that contains features extracted from 1.1 million Windows PE files using static analysis, collected in 2017. To train the model we used Random Forests, an algorithm based on binary decision trees. After some

hyper-parameter tuning we produce a model capable of recognizing malicious software that achieves around 99% accuracy at 1% false positive rate and 98.8% accuracy at 0.1% false positive rate, performing better than the EMBER proposed model. Using that model we ran a series of experiments to explore other techniques for learning. Firstly we created our own test set, containing newer malware, collected in 2018 and 2019 to observe the consequences of concept shift. The accuracy of the model dropped to 87.7% (still performing better than the proposed EMBER model) emphasizing the importance of refreshing malware detection systems with new data. Finally, we exploited retraining the model using fewer features and samples. After identifying the most useful features using information gain as a metric, we trained a new model using only 10 features instead of the 2351 used before. This new model achieves 96,1% accuracy (only 3% loss) while training takes about 1% of the time needed to train the initial model. The work is currently prepared for submission.

A study of Emotet Ransoware with Gorille, CYD Tool

Contact: Jean-Yves Marion (UL) Régis Lhoste (CYD) L. Robin (CYD)

Cyber-Detect is developing an anti-virus engine named Gorille. It is based on morphological analysis which is devised at LORIA, the computer science lab of Lorraine University. Both partners UL and CYD conducts regularly studies on ransomware in order to demonstrate the ability of Gorille to detect ransomware that may be difficult to catch by more traditional approach. The former study was about LockerGoga and the new one is about Emotet. Variants of Emotet pass under the radars of anti-viruses, because they are obfuscated. The French national organization ANSSI warned of this situation in the alert bulletin dated September 7, 2020. Emotet is a banking Trojan horse that has been around since 2014. Emotet has evolved and the new versions include new services for spreading spam and malware. According to the US Department of Homeland Security, the damage caused by Emotet would amount to $1M, making it one of the most expensive malware. We looked at a variant of EMOTET. We submitted it to Virus Total on October 14, 2020. Only seven out of the sixty-two anti-virus programs on the Virus Total platform detected the presence of malicious code in the executable. To go further, we performed a dynamic analysis, which consists in examining a sample in a sandbox and then in performing a morphological analysis. We established that the first wave of code in this sample contains an open-source code application that has been modified in order to decipher the program that will allow to install Emotet.

Gorille Cloud

Contact: Jean-Yves Marion (UL) Régis Lhoste (CYD) L. Robin (CYD), C. Jannier (CYD), E. Werner (CYD), J. Derozier (CYD), F. Sabatier (UL)

CYD, with UL, has pursued the development of Gorille. Gorille is a tool that analyzes highly obfuscated binary codes (Windows PE X86) like malware. The technology used by Cyber-Detect leans on the concept of morphological analysis developed at Lorraine University. It is based on the automatic construction of multi-dimensional signatures, stored in a knowledge behavior graph databases, and so capturing binary code behavior and functionalities.

CYD has released a first version of *Gorille Cloud*, which is accessible on demand at . We also released an API in order to access the information generated by Gorille. The API is compatible with the one of Virus ToTal V3 and the document is here : . The Gorille API allows to write scripts to upload and scan files.

Gorille Cloud is the result of an important engineering efforts of development. It embeds a several morphological analysis engines and several sand-boxes to perform dynamic analysis. As a result, Gorille Cloud should be able to answer to effectively to requests. The next step is to integrate Gorille Cloud in Concordia platforms.

Extracting Executable Payloads From Packed Malware

Contact: Jean-Yves Marion (UL) Sylvain Cecchetto (UL)

Over the past two decades, packed malware is always a veritable challenge to security analysts. Not only is determining the end of the unpacking increasingly difficult, but also advanced packers embed a variety of anti-analysis tricks to impede reverse engineering. As malware?s APIs provide rich information about malicious behavior, one common anti-analysis strategy is API obfuscation, which removes the metadata of imported APIs from malware?s PE header and complicates API name resolution from API callsites. In this way, even when security analysts obtain the unpacked code, a disassembler still fails to recognize imported API names, and the unpacked code cannot be successfully executed. Recently, generic binary unpacking has made breakthrough progress with noticeable performance improvement. However, reconstructing unpacked code?s import tables, which is vital for further mal- ware static/dynamic analyses, has largely been overlooked. Existing approaches are far from mature: they either can be easily evaded by various API obfuscation schemes (e.g., stolen code), or suffer from incomplete API coverage. In this paper, we aim to achieve the ultimate goal of Windows malware unpacking: recovering an executable malware program from the packed and obfuscated binary code. Based on the process memory when the original entry point (OEP) is reached, we develop a hardware-assisted tool, API-Xray, to reconstruct import tables. Import table reconstruction is challenging enough in its own right. Our core technique, API Micro Execution, explores all possible API callsites and exe- cutes them without knowing API argument values. At the same time, we take advantage of hardware tracing via Intel Branch Trace Store and NX bit to resolve API names and finally rebuild import tables. Compared with the previous work, API-Xray has a better resistance against various API obfuscation schemes and more coverage on

resolved Windows API names. Since July 2019, we have tested API-Xray in a production environment to assist security professionals in malware analysis: we have successfully rebuilt 155,811 executable malware programs and substantially improved the detection rate for 7, 514 unknown or new malware variants. This result is accepted at Usenix 2021 [26].

Neural Network-Based Side Channel Attacks and Countermeasures

Contact: Dimitrios SERPANOS (ISI) S. Yang, M. Wolf

In this work, we surveyed results in the use of neural networks and deep learning in two areas of hardware security: power attacks and physically-unclonable functions (PUFs). Hardware security techniques are prone to the same cat-and-mouse interplay as other forms of computer security. Neural networks are a recent addition to the toolbox of both attackers and defenders. We surveyed results in two aspects of the use of deep learning and neural networks for hardware security: power attacks and physically-unclonable functions (PUFs). the work has been published in [129]

6.6.3 Detecting Service dependencies

Contact: Yiqun Chen (ULANC)

SlowCoach: Mutating Code to Simulate Performance Bugs Contact: Yiqun Chen [ULANC]
Performance testing is a technique to identify performance bugs caused by unnecessarily inefficient codes. However, the efficacy of performance tests remains unknown with limited evaluation. Given the success of mutation testing (MT) in evaluating the effectiveness of functional tests, performance mutation testing (PMT) is proposed to simulate performance bugs in software products. A key challenge of PMT is that injecting performance bugs must not alter the functional behavior of the original program. Moreover, it is cumbersome to assert if a performance mutant should be killed if no performance impact is observed, as performance measurement could be biased or performance tests do not exercise the buggy code sufficiently. We aim to address these two aforementioned challenges for PMT and propose an easy-to-extend mutation testing framework called SlowCoach, which is able to apply complicated mutation heuristics by configuration and illustrate the number of executions of the mutated code.

6.6.4 System security validation and zero-days

Contact: Jean-Yves Marion (UL) Tristan Benoit (UL)

Binary level toolchain provenance identification with graph neural networks
Contributors: Tristan Benoit, Sebastien Bardin, Jean-Yves Marion (UL)
Contact Person: Jean-Yves Marion [UL]

Identifying the *toolchain provenance*, i.e. the compiler family (e.g. Visual Studio), the compiler version (e.g. 10.0, 12.0) and its optimization options (e.g. $-O1$, $-O2$), that have been used to produce a given bare binary code is an important problem in at least two scenarios:

- *Determination of security flaws inside binary codes.* Applications are often built by linking together commercial off-the-shelf libraries (COTS). (More than 70% of commercial applications used COTS (Gartner).) While allowing faster development cycles, developers do not have the source code of these COTS and do not know the compiling chain used to generate them. This is an important issue in software maintenance and long-term support as compilers may inject vulnerabilities that are discovered after the COTS released and after the deployment of the applications that used them. For example, CVE-2018-12886 describes a vulnerability allowing an attacker to bypass stack protection in GCC 4.1 though 8. Hence, there is a need to be able to retrieve the compiling chain in order to assess whether an application may present a certain vulnerability;
- *Identification of known functions.* Library function identification in a binary code is another primary issue for software maintenance and security, such as clone detection using Deep Learning) or malware reverse engineering. The function name identification problem is readily solved when the binary code under analyse is well-behaved, that is when it contains enough information to disassemble it.

We consider the problem of recovering the compiling chain used to generate a given bare binary code. We present a Graph Neural Network framework at the binary level to solve this problem, with the idea to take into account the shallow semantics provided by the binary code's structured control flow graph (CFG). We introduce a Graph Neural Network, called Site Neural Network (SNN), dedicated to this problem. To attain scalability at the binary level, feature extraction is simplified by forgetting almost everything in a CFG except transfer control instructions and performing a parametric graph reduction. Our experiments show that our method recovers the compiler family with a very high F1-Score of 0.9950 while the optimization level is recovered with a moderately high F1-Score of 0.7517. On the compiler version prediction task, the F1-Score is about 0.8167 excluding the clang family.

This work has been accepted at MLPA workshop [16] and submitted to SANER conference 2021.

7 Data/Application-Centric Security (T1.4)

Task 1.4 (T1.4) of the CONCORDIA project is concerned with data and application-centric security. Continuing our efforts from the first year of the project, we base our approach on the general framework proposed by TUBS (as the task leader). Figure 13 presents this framework in order to organize the research efforts in this task as well as to show the possible collaboration efforts with the different tasks in the other WPs, in a consolidated manner. Task T2.1 (Telecom pilot) and Task T2.2 (Finance pilot) have agreed to adopt this framework. Data provided by these pilots, will be processed by different analysis mechanisms to detect any Indicators of Compromise (IoC) and to determine the credibility and seriousness of a potential threat. Later, we share information related to these IoCs with the CONCORDIA central threat intelligence platform (MISP) that is provided by T3.1. Within this framework, we are putting more emphasis on cloud security to solve the next issues:

- How to protect (big) data before and after storing it on the cloud.
- How to protect the cloud services themselves.

Another critical question is how to perform behavioral analysis on the collected data to detect suspicious behaviors or attacks. Forensic data visualization is the approach we use to detect malicious activities. This approach can offer important information on suspicious patterns in a network, which can the be detected by graph analysis algorithms. Models can, subsequently, be built to prevent frauds from happening in the network.

7.1 Overview

In the following paragraphs, we report on the research activities performed within T1.4. Section 7.5.1 presents a solution for protecting cloud services using auto-

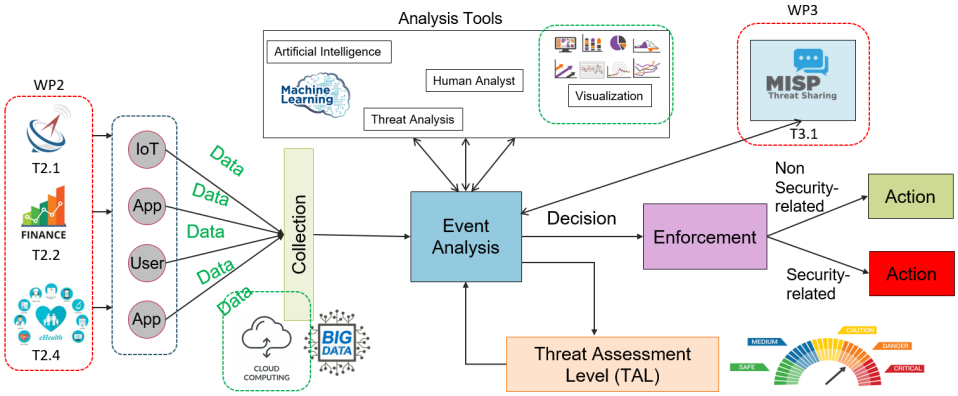

Figure 13: Framework for application/data-centric security

mated security enhancement strategies, with a particular focus on challenges related to the migration of resources composing these services. Section 7.5.2 discusses the continuous certification of composed cloud services, as well as the cost prediction of this operation. Section 7.5.3 presents a mechanism to enable the strengthening of an application against Code Reuse Attacks, based on security policies enforced by the Memory Management Unit (MMU) of the Linux kernel. Section 7.5.4 presents a policy-based model for Vehicle-to-Cloud communication, which allows to encrypt and control the access to messages published by vehicles. Section 7.5.5 presents enhancements to the previously developed framework for creating and managing cyber insurance policy for cyber-systems. Section 7.5.6 presents a solution that involves actual 5G slicing techniques to deal with limitations in the healthcare sector for elderly care. Section 7.5.7 presents a multi-layered security solution that includes an authentication mechanism and a machine learning platform to deal with attacks in the cellular IoT domain. Section 7.5.8 discusses the use of graph database mechanisms for cyber network analysis and visualization. Section 7.5.9 discusses a practical attestation system enabling seamless attestation between the hosted applications and the end clients, with low latency offered by a novel caching mechanism. Section 7.5.10 presents a scheduling approach for network packet processing applications.

7.2 Link between T1.4 work and CONCORDIA pilots

Collaboration between T1.4 and some of the project pilots has been ongoing and has yielded some interesting results. In this sense, we consider it one of the highlights of the task for the second year. Figure 14 indicates these relationships.

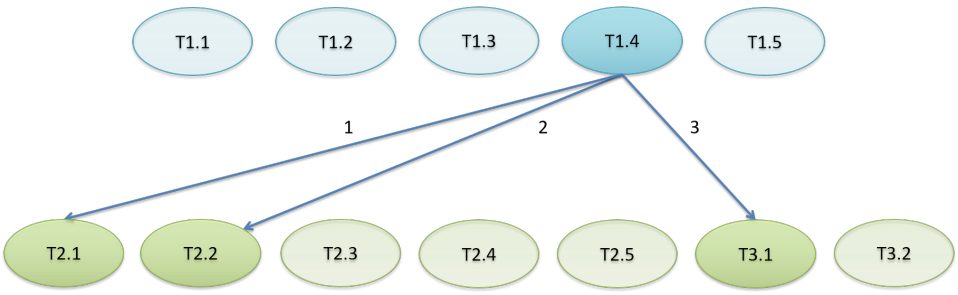

Figure 14: Links between task T1.4 and the CONCORDIA pilots

Strong collaboration has been developed with the telecommunications sector (T2.1). As WP1 deals with the scientific research within CONCORDIA which is mostly performed by the academic partners, specifically in the context of T1.4 they are able to contribute to T2.1 by researching and developing software that adheres to the requirements set by the Telecom sector partners. Derived implementations can also be integrated into the Threat Intelligence Platform, with the option to propose them for integration directly into the MISP open source project. Collaboration be-

tween the two tasks has been strengthened further, by taking part in virtual meetings where issues raised by the respective partners as well as possible solutions are being discussed.

Additionally, collaboration with the finance sector (T2.2) has been ongoing. Our proposed framework (as shown in Figure 13) has been integrated into the use case architecture and technologies proposed/developed by academic partners (secure data exchanging protocol, big data analysis based on blockchain, etc.) have been taken into consideration for the Financial Threat Intelligence Platform.

Lastly, as we have mentioned before, all the collected information about the IoCs are shared with the project's Threat Intelligence Platform (T3.1). TELENOR and TIM have established and support the collaboration between T3.1 and T1.4 partners.

7.3 Security Roadmap Considerations

In order to achieve digital sovereignty and increased levels of information technology security at the European level, it is important to identify research challenges that can act as enablers for the European industry to build the most secure products in the world (Security made in Europe). Similarly to the cases of the other tasks, we present here possible future research directions, that are specific to data/application security.

7.3.1 EU-controlled Cloud Infrastructure (GAIA-X)

The European Union aims to create GAIA-X, a secure and federated digital ecosystem that meets the highest standards of digital sovereignty, by combining existing central and decentralized infrastructures. Consequently, common requirements derived from all European partners, openness, transparency and use of secure, open technologies are important and will be used as foundations on which the framework aims to be built. It is thus necessary to provide access to secure, trustworthy and automated services and API-controlled infrastructures. Solutions must be able to minimize the leak/loss of data and increase security in software/applications development, in order to facilitate increased data value and support cross-sector cooperation.

7.3.2 Smart Technologies

The future of the facilitation of everyday life lies in smart technologies. Smart and green energy systems will generate electricity, store it and interact with the power grid in order to provide the necessary energy. Smart health monitoring systems will provide care based on distributed data and intercommunication with other systems or actors (e.g., doctors). Smart commerce will facilitate international activities based on multiple types of data as well as numerous stakeholders. Hence, it becomes increasingly necessary to provide secure and smart centralized management

of identities and access to all exchanged data and applications for each distinct domain. Best practices must include end-to-end security of an application and its communication with external services, data confidentiality / integrity / availability / anonymity, privacy controls over accessibility at different levels with respect to actors and compliance with related assurance and certification standards.

7.3.3 Securing data/software in distributed computing environments

Internet-of-Things has been on the rise and with the imminent adoption of 5G, it will continue to grow even more, creating multitudes of networks where data is being exchanged among and applications are executed on the different components. In these multi-device distributed environments, data can be used to provide integrity and trust among the communicating entities / running software, by securely identifying all involved parties. Operating systems driving such data/software, as well as the ability to securely update them, also play an important role in such environments. It is, thus, important to be able to provide solutions that secure this kind of data, their exchange and the applications that depend upon them. We expect research in the future to tackle these important subjects as well.

7.3.4 Inter-networking in the future

Data flows through the Internet in massive amounts. However, users don't usually have a say in how their data is being processed and handled: who is responsible, where it is stored, in what format, under what security measures, etc. Furthermore, data can be intentionally mishandled or even used to launch cyber attacks (DDoS, phishing, etc.). Evidently, it is important to provide security mechanisms that can assure the proper handling of data based on advertised security properties. Additionally, solutions need to provide users with the ability to verify that their data is being processed in the way they want.

7.4 Impact of COVID-19 on Research Activities

COVID-19 has not impacted the research efforts of most partners in a major way. That being said, all plans for physical meetings naturally had to be canceled. Hence, all collaboration efforts moved to online teleconferences which are held at regular intervals. While virtual meetings provide significant help in continuing collaboration among partners, understandably they cannot fully replace face-to-face meetings where the personal communication often leads to new venues to explore.

Unfortunately, however, in two specific cases, those of OsloMet and Telenor, the pandemic had some serious ramifications. Due to the lockdown rules in Oslo, Norway, these partners weren't (and still aren't) allowed to access their premises, so they aren't able to conduct the necessary research work. Furthermore, an unforeseen consequence was that their Machine Learning Platform suffered a severe malfunction which hindered even more their research activities. They are currently

in the process of purchasing new units in order to resume work, something that most likely will be possible only after the beginning of next year.

7.5 Summaries of T1.4 research efforts

The sections hereafter contain detailed summaries of the research efforts undertaken within the T1.4 context.

7.5.1 Automating Security Enhancement for Cloud Services

Contact: Remi Badonnel, Mohamed Oulaaffart, Olivier Festor (UL)

Cloud infrastructures contribute to the building of elaborated services based on multiple computing resources by composing and configuring a large variety of resources, such as virtual machines, network devices, software components. These resources may be deployed across different infrastructures owned by one or several cloud provider(s), and are subject to changes over time. This increases the complexity of management tasks and the probability of vulnerability occurrences. Such security concerns may compromise the whole cloud service. In particular, the migration of cloud resources, which consists in transferring some components of a cloud application from a given provider (or a given infrastructure) to another one, poses key security challenges.

In that context, we have pursued during this period our efforts on investigating automated security enhancement strategies for protecting cloud services. We have first analyzed and compared the properties and extensibility of cloud orchestration languages, such as the TOSCA language [105], for supporting security enhancement automation. The purpose is to take into account the horizontal and vertical dependencies that are specified by the language in the composite service description, in order to drive the security enhancement. For instance, the horizontal dependencies depend on the relationship that may exist between two interconnected resources located on different nodes, while the vertical dependencies concern a given software product with respect to the running operating system. The extended language may serve as a support for defining different orchestrated security levels, and expressing alternative configurations that permit to maintain security during migration activities. These alternatives should include the parameterization of security mechanisms (filtering rules, attack signatures, access control lists), and prevent vulnerable configurations (depending on their criticalities) by taking into account known vulnerability descriptions. They should also consider the capabilities or not of providers and resources to comply with expected modifications.

We have also started to analyze how risk management might support the selection of security mechanisms to be activated or deactivated automatically, with respect to contextual information, such as the occurrence of new threats identified by threat intelligence. The objective is to dynamically adapt the exposure of the cloud resource with respect to potential security attacks, in particular during the changes

triggered by the resource migration. The automation aims at maintaining the risk level to an acceptable level through an exposure adaptation, while minimizing the costs induced by security mechanisms [56]. For instance, when a resource is migrating to a cloud infrastructure characterized by a higher risk level, the activation of security mechanisms permits to reduce the resource exposure. In the same manner, the deactivation of security mechanisms is possible when the cloud resource is migrating to a more favorable cloud infrastructure, characterized by a lower risk level. Finally, we have worked on exogeneous security mechanisms, consisting in protecting the cloud elementary resources based on external security functions. These security mechanisms typically include firewalls, intrusion detection systems, and data leakage prevention mechanisms, and may be dynamically combined in the form of security chains by using the facilities offered by network programmability [128]. However, it is important to explore techniques to verify that these security chains are consistent and complementary to endogeneous security mechanisms. These different efforts have led to the submission of a book chapter and a conference article.

7.5.2 Security Assurance for Cloud (Composite) Services

Contact: Claudio A. Ardagna, Marco Anisetti,

The maturity reached by cloud computing has fostered the implementation of a number of distributed infrastructure, platform, and application services available worldwide. Current trends in software distribution and provisioning envision services made available as commodities over distributed systems including the Internet or the cloud marketplace. At the same time, the trend towards coarse-granularity business services, which cannot be managed by a single entity, resulted in several approaches to service composition that maximize software re-use by dynamically composing single services on the basis of their functionalities. A major challenge faced by distributed service-based systems deployed on the cloud goes beyond the ability to guarantee the functionality of composite services, and must consider the importance of guaranteeing stable Quality of Service (QoS) in the form of non-functional properties requirements such as security, performance, and trust. Service compositions need to guarantee optimal and verifiable properties, managing different events that might change their structure such as component relocation, substitution, malfunctioning, versioning, adaptation. Continuous monitoring and verification of service non-functional properties is needed and usually achieved by means of assurance techniques. Recently, certification-based assurance techniques have been introduced to guarantee stable QoS in the cloud. They are based on continuous collection of evidence on the behavior of the system, which is used to verify whether the considered system holds a specific (set of) non-functional property and award a certificate proving it. To this aim, distributed agents are instrumented to connect to different endpoints in the cloud and retrieve evidence used to evaluate the non-functional status of the target cloud-

based system. Such evidence as well as the other certification artifacts should be stored properly. In the first year of the project we extend the notion of assurance of cloud services it towards the continuous certification of composed cloud services, on one hand, and the prediction of costs for the continuous service composition certification, on the other hand. In the second year of the project we concentrate our research in improving the trustworthiness of the assurance process and in making it feasible even in hybrid cloud environments and when targeting emerging Machine Learning-based services.

We propose an approach where composite service certification meets blockchain, to support continuous and trustworthy verification of non-functional properties. The solution is enhanced by an audit process aiming to support certificates with stable properties. Trustworthiness is built on the blockchain, used as a platform for coordinating collaboration among involved parties such as service orchestrators, certification authorities, and auditors [12]. In the proposed solution the blockchain is used to ensure the trustworthiness of the orchestration and to store and partially implement the certificate verification of each component service prior to the invocation. In case the certificate verification failed the composition system can anayze the reason of the failures inspecting the evidence in the blockchain and decide to replace the service with another candidate already in cold start. To evaluate such replacement, in addition to static evaluation of certificate violations, the approach also proposes a dynamic audit that is capable to analyse past executions to evaluate the trend and preemptively substitute component services that show a decrease in their certificate quality.

Recent years have been characterized by a continuous and fast evolution of communication and computation technologies towards public infrastructures; however at the same time, the importance of private infrastructures has a comeback pushed by an increasing need of data protection, which resulted in new regulations such as the General Data Protection Regulation (GDPR) in Europe. In this complex scenario, made of hybrid systems mixing public and private infrastructures, new concerns emerged undermining the users' perceived trust, as well as their confidence in the security of the overall systems. Most of the assurance approach proposed in literature hardly cope with such hybrid systems due to the difficulties to conciliate the assurance on public cloud and the inspection on premises. In [10] we proposed a new assurance framework enabling a centralized security assurance targeting both public and private infrastructures, including public and private cloud as well as traditional private systems. It implements an assurance process that relies on a Virtual Private Network (VPN)-based solution for a smooth integration with the target system, minimizing the interferences of the framework on the target system functioning. In this paper we first defined the requirements a security assurance framework and corresponding process have to fulfill in our scenario made of hybrid systems. We then proposed a novel VPN-based assurance framework and corresponding process addressing these requirements. To this aim, we introduced several modifications to a standard VPN configuration, including

the so-called Server-side NAT, Client-side NAT, and a custom protocol to resolve conflicts between the networks in the VPN.

Modern services are increasingly substituting deterministic logic with Machine Learning (ML) models. Such models are deeply changing the services' design and development, as well as their monitoring and verification procedures. The verification of non-functional properties of software systems is a major challenge by itself and is becoming even more challenging when ML models are involved. ML models are in fact less transparent and difficult to monitor than a traditional software, thus impeding the adoption of traditional assurance approaches. In [33] we presented a discussion on a novel framework suitable for practical certification of distributed ML-powered applications in heavily regulated domains like transport, energy, healthcare, even when the certifying authority is not privy to the model training. To achieve this goal, we proposed three key ideas: *i)* use test suites to define desired non-functional properties of ML models, *ii)* use statistical monitoring of ML models' behavior at inference time to check that the desired behavioral properties are achieved, and *iii)* compose monitors' outcome within dynamic, virtual certificates for composite software applications. In [13] we presented a methodology based on Multi-Armed Bandit (MAB) for the evaluation of non-functional properties of ML models, which represents the cornerstone for a future certification of ML-based software systems. Specifically we considered a scenario where multiple ML models are available and can be selectively compared in terms of their non-functional properties. Our solution can be initially adopted at development time to select the best model to be integrated in the system (static evaluation) and then used to monitor the model behavior at run time (dynamic evaluation). In the latter case, ML models evolve over time (e.g., partial re-training/tuning) and therefore their performance in terms of non-functional property support is monitored possibly triggering substitution of the deployed model.

7.5.3 Securing Applications via the MMU, Based on Access Control Policies

Contact: Vassilis Prevelakis, Marinos Tsantekidis (TUBS)

When an adversary tries to launch a Code Reuse Attack, they identify sequences of useful instructions already present in a program's memory space and then group them together in order to trick the CPU into performing some actions not originally intended by the running program. This is due to the fact that the execution can move anywhere within a process's executable memory area, as well as the absence of policy checks when a transfer is performed.

In our effort to defend against this type of attacks, we demonstrate how we can intercept library calls using wrappers on the user side, which however works only with dynamically linked code and requires labor-intensive processing of the call arguments. Mirroring the concept of library call interception to the kernel side, we present a Proof-of-Concept mitigation technique based on a modified Linux kernel

where each library - either dynamically or statically linked - constitutes a separate code region. The idea behind this technique is to compartmentalize memory in order to control access to the different memory segments, through a $gate$. Taking our previous work one step further, we also present an updated version of our kernel-side technique, where we implement security policies in order to identify suspicious behavior and take some action accordingly.

Furthermore, we can extend our mechanism to include other aspects of memory management, namely: (a) a reachability analysis to determine all paths among all separate memory areas, (b) finer-grained segmentation of each area based on its functionality in order to further lower the chances of an attacker's success, (c) private data pages where we use the MMU to ensure that each area has private memory, which is not accessible by code running in other regions of the same process and (d) stateful execution policies that help us maintain state across $gate$ invocations.

Additionally, we can apply our approach in educational training scenarios, in order to allow students/trainees to use their imagination and creativity in exploring how an application behaves when an attack is performed.

7.5.4 SPPS: Secure Policy-based Publish/Subscribe System for V2C Communication

Contact: Vassilis Prevelakis, Mohammad Hamad (TUBS)

The Publish/Subscribe (Pub/Sub) pattern is an attractive paradigm for supporting Vehicle to Cloud (V2C) communication. However, the security threats on confidentiality, integrity, and access control of the published data challenge the adoption of the Pub/Sub model. To address that, in this work [63] we propose a secure policy-based Pub/Sub model for V2C communication, which allows to encrypt and control the access to messages published by vehicles.

A vehicle encrypts messages with a symmetric key while saving the key in distributed shares on semi-honest services, called KeyStores, using the concept of secret sharing. The security policy, generated by the same vehicle, authorizes certain cloud services to obtain the shares from the KeyStores. Here, granting access rights takes place without violating the decoupling requirement of the Pub/Sub model.

Experimental results show that, besides the end-to-end security protection, our proposed system introduces significantly less overhead (almost 70% less) than the state-of-the-art approach (SSL) when reestablishing connections, which is a common scenario in the V2C context due to unreliable network connection.

7.5.5 A Framework for Liability Based Trust

Contact: Sotiris Ioannidis (FORTH) Giorgos Christou (FORTH)

The establishment of trust across interconnected cyber systems is very important. Widely approved, effective means of managing risks and uncertainty are cyber insurance and security certification. Certification is one way to offer and establish trust relations, since it provides the necessary evidence of the required regular assessment for the provision of a service against security control measures that are explicitly designed to defend against security risks. By definition, insurance enables trust, as it (i) establishes the responsibility of covering reinstating service provision costs following after interruptions or deviations from contractual obligations and/or regulatory standards and (ii) provides compensation for losses, suffered by service consumers due to improper service provision Certification and insurance have been used as instruments for risk mitigation and trust establishment

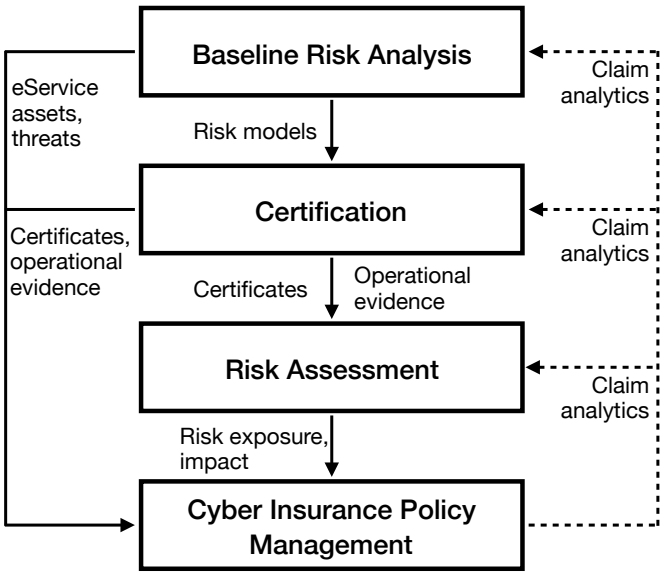

Figure 15: Generic process for cyber insurance management.

In this work, we develop CYBERSURE as shown in the figure, a framework that supports the creation and management of cyber insurance policies in order to establish trust in cyber systems and services. This framework is supported by a platform of integrated tools that enable: (i) the dynamic certification of security and privacy properties of cyber systems and services that need to be insured, (ii) the dynamic estimation of security and privacy risks for such systems and services, and finally, (iii) the development, monitoring and management of cyber insurance policies for these systems and services.

7.5.6 Secure Healthcare: 5G-enabled Network Slicing for Elderly Care

Contact: Bruno Dzogovic (OsloMet), Thanh van Do (TELENOR)

While the research and advancements in 5G progresses, the focus of its application is shifting towards the industry and different verticals. The healthcare sector is also where 5G directs its promise to accomplish the requirements for smart hospitals, assisted living, elderly care etc. In sequence of fulfilling that capability, the healthcare needs to be provided a safe, secure, reliable and robust infrastructure for handling tasks. For that purpose, in this paper we provide an insight in the existing limitations in the healthcare sector for elderly care, presenting a fashionable solution that encompasses actual 5G network slicing techniques and innovations.

7.5.7 Improving Cellular IoT Security with Identity Federation and Anomaly Detection

Contact: Bernardo Santos (OsloMet), Thanh van Do (TELENOR)

As we notice the increasing adoption of Cellular IoT solutions (smart-home, e-health, among others), there are still some security aspects that can be improved as these devices can suffer various types of attacks that can have a high-impact over our daily lives. In order to avoid this, we present a multi-front security solution that consists on a federated cross-layered authentication mechanism, as well as a machine learning platform with anomaly detection techniques for data traffic analysis as a way to study devices' behavior so it can preemptively detect attacks and minimize their impact. In this paper, we also present a proof-of-concept to illustrate the proposed solution and showcase its feasibility, as well as the discussion of future iterations that will occur for this work.

7.5.8 Forensic Data Visualization and Fraud Detection by using Graph Databases

Contact: Martina Šestak, Muhamed Turkanović (UM)

Nowadays, graph databases have found their applications in various domains where data is highly interconnected (e.g. social, computer, telecommunication and other networks). For these domains, a property graph data model can be designed to directly represent domain entities and their relationships. Modern GDBMSs such as Neo4j enable the representation of graph-oriented data and provide a declarative query syntax (e.g. Cypher) for a simpler graph data analysis. In our research, we explore how graph database technology and the available technical solutions can be used to visualize complex networks. The data visualization approach might lead to important insights on fraudulent patterns in the network, detecting the root cause of certain events, etc. We study the possibility of applying (sub)graph analysis methods and algorithms (centrality measures, clustering, etc.) for detecting anomalous patterns. These patterns can provide valuable information to build a cardinality constraint model for a given graph database. When implemented, the

model can help prevent frauds occuring in the network by limiting the number of possible relationships between a node and a subgraph.

On the other hand, in our other paper, we also studied which database solutions are used within modern blockchain platforms (e.g. HyperLedger Indie, Besu, Bitcoin, etc.). Our study revealed that the most widely used database solutions are key-value stores, i.e. LevelDB and RocksDB. We studied the properties of identified database solutions to justify their usage in the blockchain environment.

7.5.9 The Million Dollar Handshake: Secure and Attested Communications in the Cloud

Contact: Dimitris Deyannis (FORTH), Sotiris Ioannidis (FORTH)

The number of applications and services that are hosted on cloud platforms is constantly increasing. Nowadays, more and more applications are hosted as services on cloud platforms, co-existing with other services in a mutually untrusted environment. Facilities such as virtual machines, containers and encrypted communication channels aim to offer isolation between the various applications and protect sensitive user data. However, such techniques are not always able to provide a secure execution environment for sensitive applications nor they offer guarantees that data are not monitored by an honest but curious provider once they reach the cloud infrastructure.

The recent advancements of trusted execution environments within commodity processors, such as Intel SGX, provide a secure reverse sandbox, where code and data are isolated even from the underlying operating system. Moreover, Intel SGX provides a remote attestation mechanism, allowing the communicating parties to verify their identity as well as prove that code is executed on hardware-assisted software enclaves. Many approaches try to ensure code and data integrity, as well as enforce channel encryption schemes such as TLS, however, these techniques are not enough to achieve complete isolation and secure communications without hardware assistance or are not efficient in terms of performance.

Our efforts have been focused on the design and implementation of a practical attestation system that allows the service provider to offer a seamless attestation service between the hosted applications and the end clients. Furthermore, we implement a novel caching system that is capable to eliminate the latencies introduced by the remote attestation process. Our approach allows the parties to attest one another before each communication attempt, with improved performance when compared to a standard TLS handshake [25]. This works applies to the data and application-centric security in CONCORDIA.

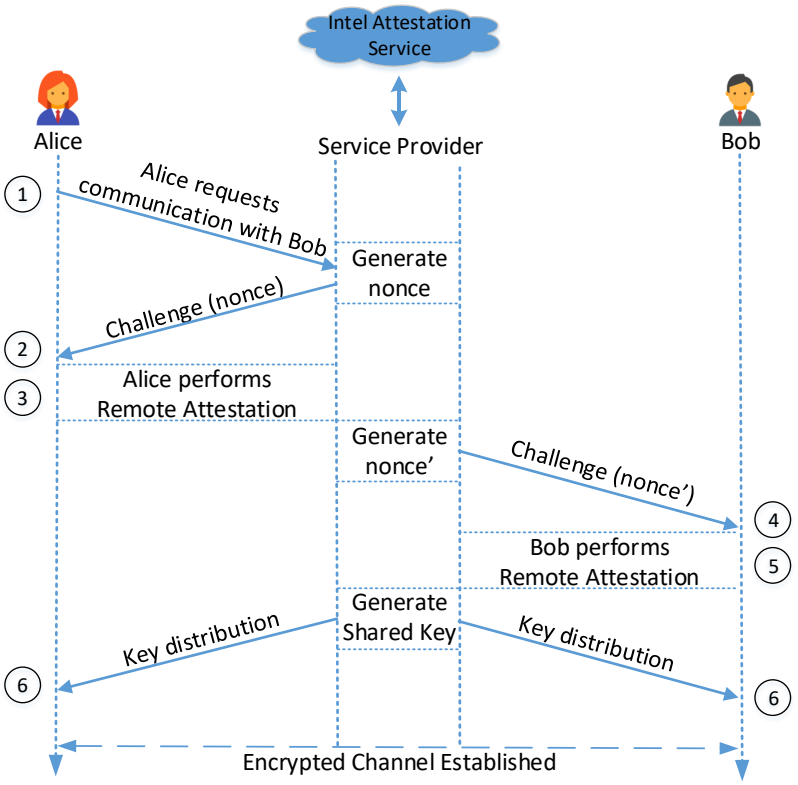

Figure 16: Attestation protocol

7.5.10 Pythia: Scheduling of concurrent network packet processing applications on heterogeneous devices

Contact: Eva Papadogiannaki (FORTH), Sotiris Ioannidis (FORTH)

Modern commodity computing systems are composed of a number of heterogeneous processing units, each one with its own unique performance and energy characteristics. However, the majority of current network packet processing frameworks targets only one device (either the CPU or an accelerator), leaving the remaining computational resources underutilized or even idle.

Our efforts have been focused on the implementation of an adaptive scheduling approach for network packet processing applications that exploits any heterogeneous architecture that can be found in a commodity high-end hardware set-up. Our scheduler not only distributes the workloads to the appropriate devices in the system to achieve the desired performance results, but also enables the multiplexing of diverse, concurrently executed network packet processing applications, eliminating the interference effects introduced at run-time. The evaluation results show that

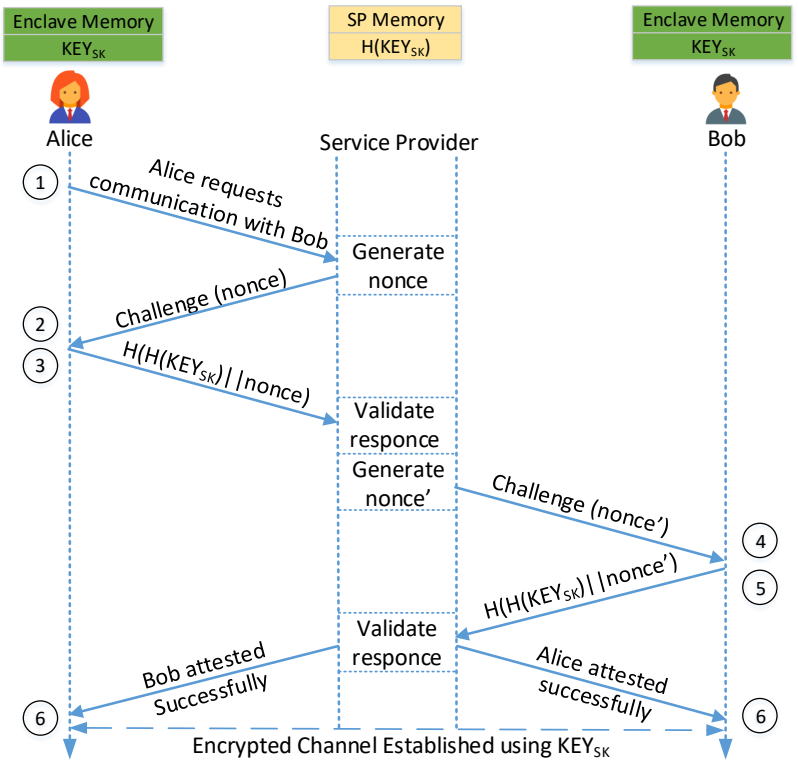

Figure 17: Attestation protocol with response caching

our scheduler is able to tackle any interference in the shared hardware resources as well to respond quickly to dynamic fluctuations (e.g., application overloads, traffic bursts, infrastructural changes, etc.) that may occur at real time [57].

In Fig. 18 we show the adaptive scheduling for different workload combinations under different conditions: network traffic rate fluctuations (a)-(b), and policy changes (c)-(d).

This works applies to the data and application-centric security in CONCORDIA.

7.5.11 Blockchain-based execution of collaborative process

Contact: Barbara Carminati (UI), Elenea Ferrari (UI)

During Y2, we have continued our research aiming at exploiting blockchain to manage a secure execution of collaborative process among organizations. Indeed, thanks to its design and consensus algorithm, blockchain provides a trustworthy infrastructure that allows partners involved in the collaboration to monitor and perform audits on the workflow transitions. In general, the focus of the existing

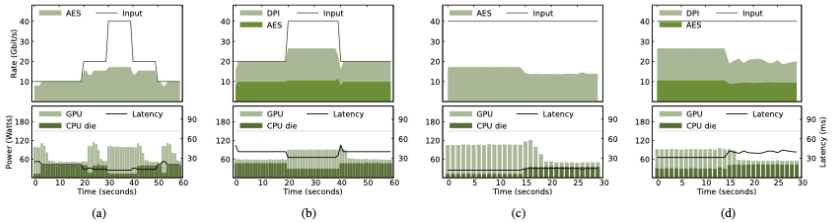

Figure 18: Adaptive scheduling for different workload combinations

blockchain-based workflow management solutions is mainly the workflow coordination. However, a challenging characteristic of some workflows is that they require the exchange of a big amount of data that has to be managed off-chain, that is, directly exchanged between data producer and consumer. This off-chain data sharing should be secured and controlled such to follow the workflow execution. To cope with this challenge, in [122] have investigated a controlled information sharing in inter-organizational workflows enforced via smart contracts. Moreover, in collaboration with UMIL, in [12] we have further exploited blockchain to design a service composition modelling collaborative approach via blockchain, driven by certified properties.

7.5.12 Access Control Policy Evaluation in NoSQL systems

Contact: Pietro Colombo (UI), Elena Ferrari (UI)

During Y2, we worked on the definition of a general approach to evaluate the effects of access control policies of the major discretionary access control (DAC) models on schemaless data stored in NoSQL systems[30]. As a matter of fact, due to the variety of data models, access control models, and related configuration options, it can be really hard for security administrators to understand the effects of access control policies on the data resources handled by their systems. The framework presented in [30] aims at mitigating this issue. The evaluation is achieved for a set of access control policies and access control configuration options by deriving a view of the protected resources that points out authorized and unauthorized contents. The approach targets schemaless data resources handled by multiple families of NoSQL systems and could even be extended to traditional DBMSs. The analysis is built on top of MapReduce, a widely popular computing paradigm that is nowadays supported by the majority of NoSQL systems. Experimental results show the approach efficiency and scalability.

8 User-Centric Security (T1.5)

Task 1.5 (T1.5) of the CONCORDIA project is focused on user–centric security. Specifically, the main research pillars of T1.5 are as follows:

- **Privacy:** This task aims on developing techniques for Privacy-Preserving Machine Learning modeling, as well as Personal Identifiable Information (PII) leakage detection to the advertising ecosystem and the general Web.
- **Identity Management:** The objective of this task is the developing blockchain based methods for creating digital identities. This will allow users of Online Social Networks (OSN) to verify their real-world identities without a centralized authority for storing and managing their personal information.
- **Social Networks and Fake News:** This task focuses on investigating techniques for identifying fake news in Online Social Networks as well as developing blockchain based methods for suppression of fake News.

8.1 Overview

During the second year of the project, the research activities of the partners resulted to fourteen publications along with one technical report and five *arxiv* preprints. One paper received Best paper nomination and another one Honorable mention in ACM conferences. Strong collaboration has been developed between academia as well as academia and industry. Specifically, UMIL and UM are working together on developing users' privacy protection solutions, TID and FORTH investigated the online tracking ecosystem and finally TID worked together with CUT on detecting aggression on Twitter and inappropriate videos on YouTube. In the following sections we present the results of the research activities related to Task 1.5 three main objectives.

Privacy: Section 8.4.1 discuses a large–scale study of mobile-specific HTML5 WebAPI calls across 183K of the most popular websites. Towards this goal, a novel testing infrastructure has been proposed which is consisting of actual smartphones on top of a dynamic Android app analysis framework so that an end-to-end exploration can be feasible. Section 8.4.2 presents an empirical evaluation of various implementations of differential privacy (DP), and measured their ability to fend off real–world privacy attacks, in addition to measuring their core goal of providing accurate classifications. This work received a Best paper nomination in ACM CCSW 2020. Section 8.4.3 discuses the issues related to Federated Learning (FL). This action presents Federated Learning as a Service (FLaaS) which can be deployed in different operational environments. FLaaS is a system enabling different scenarios of 3rd–party application collaborative model building and addressing the consequent challenges of permission and privacy management, usability, and hierarchical model training. Section 8.4.4 presents an analysis of tracking activity evolution as it is performed by websites on EU–based visitors, during a time pe-

riod of 2 years. Moreover, the connectivity of trackers with each other has been examined – specifically asks how this relates to potential cookie synchronization activity. Section 8.4.5 presents an analysis on websites with hyper–partisan, left or right-leaning focus and their online tracking practices which are imposed on their visitors. For this purpose, a methodology to systematically probe such websites was proposed in order to measure differences in user tracking. Section 8.4.6 focuses on investigating the privacy leakage (in terms of sensitive user data) in modern smartphones. In addition, work has been conducted on how mobile sensor data can be used for a plethora of attacks using the HTML5 WebAPI. Section 8.4.7 dis-cuses two research works related to privacy issues that arise in publishing knowl-edge graphs (KGs). Up to now, the protection models that have been proposed are unable to protect the users in KGs. For this, the k-Attribute Degree (k-ad) is introduced along with the k^w-Time-Varying Attribute Degree (k^w-tad) protection model in order to protect users' identities in anonymized KGs.

Identity management: Section 8.4.8 presents contribution on Privacy and access control features to the novel service PayID which has privacy and security issues. Hence, an extended version of the PayID server has been proposed which has novel features of Access Control List (ACL) and Decentralized IDentity (DID). Section 8.4.9 summarizes two research works. The first one is a machine learning method for automated signing of blockchain transactions which increases the security of the user/owner of digital resources and at the same time simplifies the digital signing process. The second one is an architecture reference model which integrates Qualified Electronic Signatures with blockchain transactions in order for enterprises and public services to leverage the blockchain technology. Section 8.4.10 presents under development work related to user-centric domain. An extended notion of behavioral assurance has been proposed in order to cope with other user behaviors in IoT context, which is the most exposed environment in terms of risks of misbehaviors. In addition, work has been conducted on users' privacy protection by developing new privacy–aware security approaches with ver-ifiable behavior. Finally, Section 8.4.11 presents Provotum and Proverum, systems for Remote Electronic and Remote Postal Voting.

Social Networks and Fake news: A series of research activities were related to social network analysis with special focus on Twitter and Youtube. Section 8.4.12 presents the findings of an analysis of 17.5M tweets (from 3M users) related to the 2020 US presidential election. The evolution of retweet graph has been studied along with a sentiment analysis on the YouTube links that were embedded in the tweets. Section 8.4.13 discusses the findings regarding the impact of State–sponsored troll accounts activities on the virality of the political information that was had been shared on Twitter during the 2016 US Presidential election. Section 8.4.14 is related to Twitter botnets discovery and classification. To this purpose, analysis has been conducted regarding the behavior and interactions of bots together with user communities. The dataset is a subset of Twitter traffic, consisting of nearly all interactions by Greek-speaking Twitter users for a period of 36

months. Section 8.4.15 presents a novel real–time framework for detecting aggression on Twitter which is based on the streaming machine learning methodology. The framework is general enough to detect other related behaviors such as sarcasm, racism, and sexism in real time. Section 8.4.16 discuses the findings related to Youtube algorithmic recommendation system which often suggests inappropriate content/videos for children. To this purpose, a classifier was build which is able to discern inappropriate content. This work received Honorable mention in AAAI ICWSM 2020. Section 8.4.17 investigates the Incel community on YouTube and in particular the evolution of this community over the last decade. Also, YouTube recommendation algorithm has been analysed regarding to what extent the recommendation drives users towards Incel-related videos. Section 8.4.18 presents an analysis of pseudoscientific misinformation and conspiracy theories on YouTube. Specifically, the users exposure to this content has been qualified in various parts of the platform along with the evolution of this exposure, based on the users' watch history. Finally, Section 8.4.19 presents an ongoing work which is related to the identification of disinformation content in Twitter. As a first step, a large Twitter dataset has been collected consisting of 850M tweets associated with COVID–19 pandemic.

8.1.1 Link between T1.5 work and CONCORDIA pilots

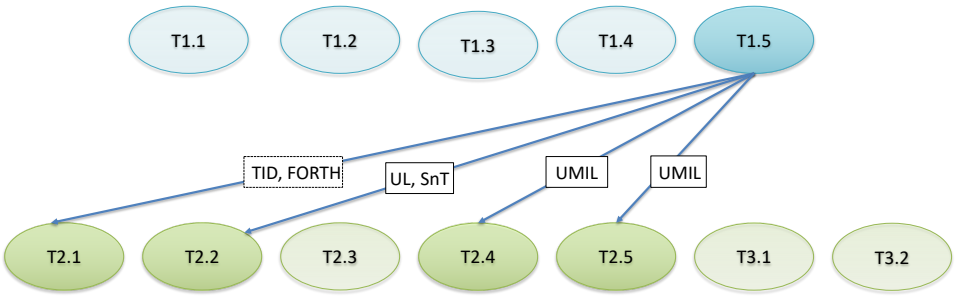

Figure 19: Links between T1.5 and the CONCORDIA pilots

The research performed within Task 1.5 is linked to the CONCORDIA pilots as follows (see Figure 19):

- The research activities of TID and FORTH which are presented in Sections 8.4.2, 8.4.3, 8.4.4 and 8.4.5 along with the relevant publications [147, 82, 134], are related to pilot T2.1.
- The research work of UL and SnT which is presented in Section 8.4.8 is related to pilot T2.2.
- The research work 8.4.10 of UMIL is related to pilots T2.3 and T2.4.

8.2 Security Roadmap Considerations

8.2.1 Fighting disinformation in Europe

Online social networks and online media platforms enable individuals from remote corners of the globe to share ideas, news, and opinions in an almost instantaneous manner. Social networks such as Twitter and Facebook have become a primary source of information for billions of users and the media where new cultural and political movements are formed and promoted. This high level of reliance on social media opened the field to malicious actors to pose new kinds of threats,which can have severe consequences at a societal level. Disinformation diffusion in social networks is one such threat carried out by diverse users who have various motives.For example, terrorist organizations deliberately diffuse false information for propaganda purposes, trying to inflict conflict or to cause extreme emotional reactions. Foreign interference of actors with motives against the EU using human or automated operated accounts (bots) can slander a candidate, trying to shift the outcome of national elections or impede the policy-making process in general.

Challenges

- **Understanding the disinformation diffusion: Multiplatform diffusion** The mechanism, the channels, and dynamics of disinformation diffusion are neither clear nor easily assessable for analysis. The disinformation content can become viral following a complex path of transmission and through many online communication platforms. The disinformation content could first be originating in the ”periphery” of social platforms and become viral in mainstream media. QAnon conspiracy theory is such an example. It is a unified–conspiracy theory consisting of several other conspiracy theories such as Pizzagate. It originated on 4chan (by the anonymous user ”Q”)and then spread through multiple social media platforms.
- **Official malicious actors: Elected politicians** Often, there is a symbiotic relationship between elected politicians and conspiracy theory promoters.Often, political parties are the source of disinformation – using as a tool the conspiracy theories aiming to create a political polarization which will consequently lead to a loyal political base. Hence, individuals who support reactionary and anti–scientific narratives can become part of the elected government. Although this is a mainly political challenge for the European democratic system, countermeasures against disinformation campaigns employed by the social platforms themselves could suppress political extremism.

Actions

- **Early detection of disinformation:Classify the content and identify the actors** One of the main challenges is detecting disinformation and misinformation operations at an early stage – before becoming viral in the main-

stream media. Therefore, research should be conducted on developing novel machine learning techniques that will classify the spread of information and identify the source of disinformation – the influential users who were responsible for the information diffusion.

- **Countering disinformation:** During crises such as the COVID–19 pandemic, false information such as pseudo scientific conspiracy theories can result in wide-spread panic and chaos. Hence, not only early detection but also countering the disinformation is crucially important. Conspiracy theories related to the origin of COVID-19 and the anti-vaccine movements could play a negative role in the fight against the pandemic. Therefore, it is cru-cial to develop countermeasures against conspiracy theories that will be, at the same time, in line with the democratic values of Europe, such as the freedom of speech. Research on the early identification of malicious users that lead to their suspension from the social platforms is one such direction. Also, it is not enough to suspend accounts spreading disinformation. It is of paramount importance to research social media dissemination strategies that increase the influence of correct fact-checking information by employ-ing graph-theoretical, game-theoretical, and human factor principles.
- **Coordination: European disinformation observatories** An integrated or federated European observatory of disinformation that will monitor the social media streams and disclose disinformation activities should be a long-term objective. The observatories are currently being established in any European country to form an internal interconnected network of national institutions. Each network hub collaborates with national authorities, fact-checking organizations, and research institutions. Research on how to properly share and aggregate information from multiple observatories could prove highly beneficial in the observatory integration effort.

8.2.2 Data Ownership and Data Privacy

The initial design requirements of the Internet and the Web in the early 60s and 90s were far different than those of today (i.e., connecting servers between academia, sharing content through simple websites, email exchange, etc.). Today, both the Internet and the Web have managed to exhibit tremendous evolvability and extendibility. They have succeeded in supporting services (e–commerce, e-banking, content distribution, video streaming, Web conferencing, etc.) and capabilities (broadband connection, mobility, satellite, etc.) that could hardly be imagined.

Online advertising and marketing appeared soon after the Web's appearance in the 90s and grew into an entire industry that is currently funding a large part of the so-called free services of the Internet. Advanced versions of web advertising and recommendation systems, in general, are heavily based on detailed personal data collected online from millions of individuals to offer tailored ad impressions and

recommendations to maximize profits of the so-called "Tech Companies," such as Google, etc. Of course, the uncontrolled user tracking and personal data collection of individuals lead to data protection and privacy problems that have challenged the Internet and the Web today.

New research efforts are required to mitigate and control the challenges mentioned above. Below we identify different directions that we need to turn our attention:

Data protection regulations: In recent years, we have witnessed new data protection regulations such as the General Data Protection Regulation (GDPR) in Europe and the California Consumers Act in the US, to name some. Since new regulations are now in place, the challenge now is shifted towards how we can apply them in practice by proactively monitoring and detecting violations in an automated way. As a result, new tools and methodologies need to be implemented to automate such regulations' enforcement. Some examples include tools related to web tracking and personal data leakage detection, website classification to identify sensitive content websites as defined by GDPR and similar legislation, Cookie consent (opt-out) automation and monitoring, browser fingerprinting mitigation, personal data handling, storage, and localization monitoring, etc.

Personal data ownership: New research needs to be conducted to allow users to have full control of their data, including their browsing patterns, shopping activities, social network activities, etc. The main focus of such tools should be but not limited to the following functionalities:

- Data portability: Data owners should be able to move their data across different online services of their choice (i.e., move financial data from one online banking service to another). As a result, new research should be focusing on novel portable data structures and mechanisms to allow the above functionality.
- Right to be forgotten: Data owners should be able to block access and delete their personal data across different online services (i.e., remove their data from a social network). New tools and methodologies need to be invented to ensure that personal data collected and stored online are under the full control of the data owner (users), rather than the data collector (online service), which is the current state that we are facing today.
- Furthermore, we need to provide technologies and tools to allow users to benefit from their personal data (i.e., create new monetization schemes based on personal data sharing).

Personal data value and Human-Centric Data economy: Most online services utilize personal data to increase their profits. For example, e-commerce websites can use personal data to train machine learning algorithms to optimize their inventory and product recommendations. The ad industry uses personal data at a massive scale to serve targeted and re-targeted advertisements at a higher premium, etc. In

all the above scenarios, the data producer (user) is only compensated by getting access to the corresponding online service for free in exchange for being tracked. Instead,it would be fairer for end-users to have direct financial benefits for their data. To provide economic benefits based on personal data, the following research questions need to be answered: What is the actual value of personal data? How can we estimate such value? What factors influence data value based on how data consumers use them? Based on what frameworks do the data owner and data user value them?

Personal Information Management Systems (PIMS): A more recent trend towards addressing privacy and cybersecurity threats around personal data is introducing an additional entity between online services and end-users. The so-called Personal Information Management Systems (PIMS) or Data Vaults. Towards that direction, we need to investigate different paradigms, such as centralized vs. decentralized PIMS, distributed open source or centralized closed source approach, and what the pros and cons of each paradigm are to achieve adaptability and global acceptance. In addition, we need to identify what the critical parts of such an ambitious approach are (i.e., data integrity, trust between nodes, data access control, etc.)

8.3 Impact of COVID-19 on Research Activities

Generally, the COVID–19 pandemic did not have a major impact on the progress of T1.5 and the partners' research activities. Obviously, all the planned meetings had to be canceled and the collaboration efforts moved to online teleconference meetings. Even though, the online meetings proved to be a very helpful tool, nevertheless they cannot replace the physical meetings and the personal communication between the partners. For this reason and in order to further support the collaboration efforts among the partners, we organized a PhD and Industry symposium among the PhD students and industry senior researchers, working on T1.5. In this event, the PhD students presented their research plan and they had the opportunity to discuss aspects of their work with senior researchers.

Unfortunately, in some partners the pandemic had some negative impact on their research. Specifically, the research to be performed by Telefonica I+D (TID) was effected by COVID-19, since it was not possible to perform the regular research internships usually performed throughout the year, due to difficulty or impossibility of traveling, inaccessibility to laboratory and other resources, etc. These had a negative impact on the productivity of the research team which was forced to scale down its research output and focus on investigation contributing to core technologies needed for Telefonica and the privacy pillar of task T1.5.

8.4 Summaries of T1.5 research efforts

In the next sections we present summaries of T1.5-related research efforts.

8.4.1 A Large-scale Study on the Risks of Mobile Sensor-based Attacks

Contact: Michalis Diamantaris (FORTH), Sotiris Ioannidis (FORTH)

Modern smartphone sensors can be leveraged for providing novel functionality and greatly improving the user experience. However, sensor data can be misused by privacy-invasive or malicious entities. Additionally, a wide range of other attacks that use mobile sensor data have been demonstrated; while those attacks have typically relied on users installing malicious apps, browsers have eliminated that constraint with the deployment of HTML5 WebAPI.

Our efforts have been focused on the implementation of a comprehensive evaluation of the multifaceted threat that mobile web browsing poses to users by conducting a large-scale study of mobile-specific HTML5 WebAPI calls across more than 183K of the most popular websites. We build a novel testing infrastructure consisting of actual smartphones on top of a dynamic Android app analysis framework, allowing us to conduct an end-to-end exploration. In detail, our system intercepts and tracks data access in real time, from the WebAPI JavaScript calls down to the Android system calls. Our study reveals the extent to which websites are actively leveraging the WebAPI for collecting sensor data, with 2.89% of websites accessing at least one sensor. To provide a comprehensive assessment of the risks of this emerging practice, we create a taxonomy of sensor-based attacks from prior studies and present an in-depth analysis by framing our collected data within that taxonomy. In Fig. 20 we present our taxonomy that aims to highlight the variety of attacks enabled by sensor data, while simultaneously obscuring the type of sensor used for each attack.

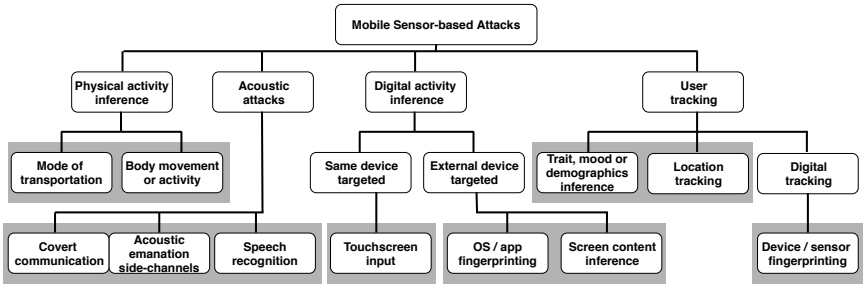

Figure 20: Taxonomy of attacks demonstrated in prior studies that leverage data from mobile sensors

We find that 1.63% of websites can carry out at least one attack and emphasize the need for a standardized policy across all browsers and the ability for users to control what sensor data each website can access [40]. This works applies to the data and user-centric security in CONCORDIA.

8.4.2 Utility vs. Privacy Tradeoffs in Differentially Private Machine Learning

Contact: Benjamin Zi Hao Zhao (University of New South Wales and Data61 CSIRO) Mohamed Ali Kaafar (Macquarie University) Nicolas Kourtellis (Telefonica)

Data holders are increasingly seeking to protect their user's privacy, whilst still maximizing their ability to produce machine models with high quality predictions. In this work, we empirically evaluate various implementations of differential privacy (DP), and measure their ability to fend off real-world privacy attacks, in addition to measuring their core goal of providing accurate classifications. We establish an evaluation framework to ensure each of these implementations are fairly evaluated. Our selection of DP implementations add DP noise at different positions within the framework, either at the point of data collection/release, during updates while training of the model, or after training by perturbing learned model parameters. We evaluate each implementation across a range of privacy budgets, and datasets, each implementation providing the same mathematical privacy guarantees. By measuring the models' resistance to real world attacks of membership and attribute inference, and their classification accuracy. we determine which implementations provide the most desirable tradeoff between privacy and utility. We found that the number of classes of a given dataset is unlikely to influence where the privacy and utility tradeoff occurs. Additionally, in the scenario that high privacy constraints are required, perturbing input training data does not trade off as much utility, as compared to noise added later in the ML process.

We presented this work in ACM CCSW [147], and it was well received, with a Best paper nomination.

This work is generally related to the first pillar of T1.5, regarding privacy-preserving machine learning.

This work is related to WP2 and telco pilot in T2.1. The collaboration is between academia and TID.

8.4.3 FLaaS: Federated Learning as a Service

Contact: Nicolas Kourtellis (Telefonica) Kleomenis Katevas (Telefonica) Diego Perino (Telefonica)

Federated Learning (FL) is emerging as a promising technology to build machine learning models in a decentralized, privacy-preserving fashion. Indeed, FL enables local training on user devices, avoid- ing user data to be transferred to centralized servers, and can be enhanced with differential privacy mechanisms. Although FL has been recently deployed in real systems, the possibility of collaborative modeling across different 3rd-party applications has not yet been explored. In this paper, we tackle this problem and present Federated Learning as a Service (FLaaS), a

system enabling differ- ent scenarios of 3rd-party application collaborative model building and addressing the consequent challenges of permission and privacy management, usability, and hierarchical model training. FLaaS can be deployed in different operational environments. As a proof of concept, we implement it on a mobile phone setting and discuss practical implications of results on simulated and real devices with respect to on-device training CPU cost, memory footprint and power consumed per F L model round. Therefore, we demonstrate FLaaS's feasibility in building unique or joint F L models across applications for image object detection in a few hours, across 100 devices.

We are presenting this work in ACM Distributed ML [82].

This work is generally related to the first pillar of T1.5, regarding privacy-preserving machine learning.

This work is related to WP2 and telco pilot in T2.1. The collaboration is between academia and TID.

8.4.4 Evolution of the Online Tracking Ecosystem

Contact: Konstantinos Solomos (University of Illinois at Chicago) Panagiotis Ilia (University of Illinois at Chicago) Sotiris Ioannidis (FORTH) Nicolas Kourtellis (Telefonica)

Websites are constantly adapting the methods used, and intensity with which they track online visitors. However, the wide-range enforcement of GDPR as of May 2018 forced websites serving EU-based, online visitors to eliminate or at least reduce such tracking activity, given they receive proper user consent. Therefore, it is important to record and analyze the evolution of this tracking activity and assess the overall" privacy health" of the Web ecosystem and if it is better after GDPR enforcement. This work makes a significant step towards this direction. In this paper, we analyze the online ecosystem of 3rd-parties embedded in top websites which amass the majority of online tracking through 6 time snapshots taken every few months apart, in the duration of the last 2 years. We perform this analysis in three ways: 1) by looking into the network activity that 3rd-parties impose on each publisher hosting them, 2) by constructing a bipartite graph of" publisher-to-tracker", connecting 3rd parties with their publishers, 3) by constructing a" tracker-to-tracker" graph connecting 3rd-parties who are commonly found in publishers. We record significant changes through time in number of trackers, traffic induced in publishers (incoming vs. outgoing), embeddedness of trackers in publishers, popularity and mixture of trackers across publishers. We also report how such measures compare with the ranking of publishers based on Alexa. On the last level of our analysis, we dig deeper and look into the connectivity of trackers with each other and how this relates to potential cookie synchronization activity.

We presented this work in IFIP / IEEE TMA [134].

This work is generally related to the first pillar of T1.5, regarding PII leakage and tracking detection on the Web.

This work is related to WP2 and telco pilot in T2.1. The collaboration is between academia and TID.

8.4.5 Differential Tracking of Users on Hyperpartisan websites

Contact: Pushkal Agarwal (King's College London) Sagar Joglekar (King's College London) Panagiotis Papadopoulos (Telefonica) Nishanth Sastry (King's College London) Nicolas Kourtellis (Telefonica)

Websites with hyper-partisan, left or right-leaning focus offer content that is typically biased towards the expectations of their target audience. Such content often polarizes users, who are repeatedly primed to specific (extreme) content, usually reflecting hard party lines on political and socio-economic topics. Though this polarization has been extensively studied with respect to content, it is still unknown how it associates with the online tracking experienced by browsing users, especially when they exhibit certain demographic characteristics. For example, it is unclear how such websites enable the ad-ecosystem to track users based on their gender or age.

We presented this work in ACM The Web Conference [5].

This work is generally related to the first pillar of T1.5, regarding PII leakage and tracking detection on the Web.

This work is related to WP2 and telco pilot in T2.1. The collaboration is between academia and TID.

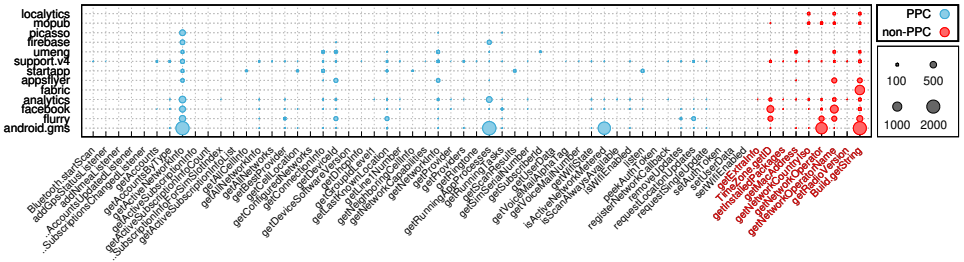

Figure 21: PII leakage from the most popular third-party libraries

8.4.6 Privacy leakage in modern smart phones

Contact: Michalis Diamantaris (FORTH), Sotiris Ioannidis (FORTH)

FORTH is currently focusing on investigating the privacy leakage in modern smartphones since their ubiquitous nature has rendered them a treasure trove of sensitive user data and personally identifiable information. Our previous study [41] on over

5K popular apps demonstrates the large extent to which personally identifiable information is being accessed by libraries and highlights the privacy risks that users face. We find that an impressive 65% of the permissions requested do not originate from the core app but are issued by linked third-party libraries, 37.3% of which are used for functionality related to ads, tracking, and analytics. In Fig. 21 we present personally identifiable information (PII) leakage from the most popular third-party libraries (sorted in descending order) broken down to the corresponding function call used.

Blue circles denote PII being accessed through permission-protected calls, while the red circles indicate PII access by functions that are not permission-protected. The size of the circle denotes the number of apps in each case. We have also worked on how mobile sensor data can be used for a plethora of attacks using the HTML5 WebAPI. Currently, our research focuses on identifying security and privacy threats that mobile users face and designing protection mechanisms for better protecting users. This works applies to the data and user-centric security in CONCORDIA.

8.4.7 Anonymization of Knowledge Graphs

Contact: Barbara Carminati (Universita' degli Studi dell'Insubria) (UI) Elenea Ferrari (Universita' degli Studi dell'Insubria) (UI) Anh-Tu Hoang (Universita' degli Studi dell'Insubria) (UI)

During Y2, we have continued to investigate the privacy issues that might arise in publishing knowledge graphs. Indeed, more and more data providers share users' data by using knowledge graphs (KGs) as these graphs can represent many types of users' attributes and relationships. Although many protection models have been presented to protect users in anonymized data, they are unsuitable to protect the users in KGs. To cope with this problem, since KGs are directed graphs, we first focused on previously defined protection models tailored for these graphs, namely the Paired k-degree[24] and the K-In&Out-Degree Anonymity [146]. While these models are sound, the proposed anonymization algorithms cannot always generate anonymized graphs. Thus, in [68], as a first step, we proposed a more reliable anonymization algorithm, i.e., the Cluster-Based Directed Graph Anonymization Algorithm (CDGA), to generate anonymized directed graphs satisfying the requirements of both the Paired k-degree and the K-In&Out-Degree Anonymity. After dealing with directed graphs, we shift our research to anonymize KGs containing both types of users' information: attributes' values and relationships. Adversaries can re-identify users in anonymized KGs by combining both information types of their victims. Unfortunately, previously mentioned protection models, i.e., the Paired k-degree[24], and the K-In&Out-Degree Anonymity [146], cannot be applied as they only anonymize one relationship type of these users. To overcome this limit, in [69], we proposed k-Attribute Degree (k-ad), an extension of k-anonymity, the Paired k-degree [24], and the K-In&Out-Degree Anonymity[146], to protect

users' identities in anonymized KGs. Moreover, we designed the Cluster-Based Knowledge Graph Anonymization Algorithm (CKGA) to anonymize KGs, according to the proposed k-Attribute Degree (k-ad) protection model. To allow for more flexibility, we aimed at allowing data providers to specify which clustering algorithm they want to use to generate clusters. To reach this goal, our algorithm performs an additional step that generates data points for users in the given KG such that the Euclidean distance of two points is almost equal to the information loss of making the attributes' values and out-/in-degrees of users corresponding to these points identical. Since most state-of-the-art clustering algorithms (e.g., k-means, HDBSCAN) can take as input these points, our algorithm allows data providers to use most clustering algorithms to generate clusters. To prevent the generated clusters from having less than k users, we proposed the k-Means Partition Algorithm (KP) to ensure that all of the generated clusters have from k to $2 \times k - 1$ users. The experiments, carried out on five real-life datasets, showed that CKGA outperforms previous algorithms (i.e., [68, 24, 146]). Although k-ad prevents adversaries from re-identifying users in an anonymized KG, they can still re-identify these users if they exploit many versions of the anonymized KG. Therefore, in [70], we developed the k^w-Time-Varying Attribute Degree (k^w-tad) protection model, to protect identities of users appearing at least once in w continuous anonymized KGs, where w and k are two positive numbers provided by data providers. The providers can use w to control how many continuous anonymized KGs to monitor and k to control the confidence of re-identifying their users in these KGs.

8.4.8 KYC, Decentralized Identity, Privacy-Preserving PayID

Contact: Wazen Shbair (University of Luxembourg) (UL)

Our contribution is twofold, first we continue our collaboration with the industrial partner in exploring the advantage of blockchain-based KYC approaches. As next step, we plan to deploy the solution developed by our team in their premises for real testing and assessing the integration process. The second contribution is related to Privacy and access control features to the novel service PayID released by Ripple [4]. PayID is an international to replace all hard to remember cryptocurrency addresses with a human-readable identifier for all payments. Complicated payment addresses presents a major barrier to adoption of blockchain and other payments technology. We have assessed the reference implementation of PayID and found it has privacy and security issues. We proposed an extended version of the PayID server with novel features of Access Control List (ACL) and Decentralized IDentity (DID). In the context of this task we found that our contribution may solve the issue of impersonation issue raised by the financial industry partner. Thus we have deeply explored the decentralized identity solutions with focus on blockchain-based one that comes from Hyperledger Indy framework. Our contribution is linked to T2.2 the financial pilot.

8.4.9 Blockchain transactions and digital signing processes

Contact: Blaž Podgorelec (UM), Muhamed Turkanović (UM)

Blockchain technology allows users to transfer digital resources ownership (e.g., cryptocurrency) directly to each other within a decentralized and distributed peer-to-peer network, without the need for third-party (e.g., financial institution). The base of blockchain data, stored in distributed ledgers, are digitally signed transactions. Data can be appended in the ledger only after a user with a digital identity under his control perform a digital signing process of the blockchain transaction. But, the process of digital signing is time-consuming and not user-friendly, which may be a reason why blockchain technology is even today not fully accepted by users and enterprises. Moreover, studies on the usability of digital signatures report that there are still many barriers preventing users from accepting digital signatures within their everyday usage. These barriers are associated with managing, controlling, and indeed using cryptographic keys. Furthermore, studies have pointed out that these usability issues also impact users' security, e.g., users cannot recognize potential intrusions while being involved in a digital signing process. To tackle those above-to-user-centric security-related issues, we have developed and proposed a machine learning-based method that introduces automated signing of blockchain transactions while including a personalized identification of anomalous transactions [112]. As a result, the security of users owned digital resources has increased, and at the same time, the digital signing of blockchain transactions for the user is simplified. Still, digital transactions within the blockchain network are signed by blockchain-based digital identity. Simultaneously, blockchain technology's acceptance is increasingly among private enterprises and public services (e.g., EBSI - European Blockchain Services Infrastructure). In those environments, where the included blockchain network is usually permissioned or public-permissioned, challenges related to aligning blockchain identity management schemes with the Public Key Infrastructure (PKI) and the Qualified Digital Certificates issued by Qualified Trust Service Providers appears. To solve this challenge, we introduced a solution [143] in the form of an architecture reference model, enabling enterprises and public services to leverage blockchain technology by integrating Qualified Electronic Signatures with blockchain transactions. The architecture reference model's evaluation was provided by designing a Blockchain-based Trusted Public Service and a use-case scenario example. The proposed architecture reference model is based on the existing CEF building blocks such as EBSI, eSignature, and eID compliant with eIDAS.

Our work results that were briefly described above results in improved users' security, mostly, but not exclusively limited to blockchain-based environments.

8.4.10 Assessing Trust Assurance based on Human Behavior Compliance

Contact: Marco Anisetti Claudio A. Ardagna

The success of the Internet of Things (IoT) is increasingly pushing towards the development of Cyber–Physical Systems (CPSs) that go beyond traditional IT boundaries, combining physical and digital environments in a single one. Today systems have an opaque perimeter, where a mixture of platforms, software, services, things, and people collaborate in an opportunistic way. Computations are then moving from the center (e.g., cloud) to the periphery, supporting analytics and knowledge extraction partially executed at the edge of the network, near the physical environment and sensors where data are collected. In this scenario, people carry/manage a plethora of smart devices (often integrated in their smartphones or in their homes) sensing the surrounding environment and communicating with edge nodes without even noticing. Like for the Web 2.0 revolution when user-generated content entered the loop, people are not only passive consumers of pervasive services; they rather become (often unintentional) service providers that distribute user-collected data. The exponential growth of smart devices (200 billions of connected objects with a mean of 26 objects for every human predicted by 2020) and connected people (2.87 billion users carrying a smartphone by 2020), coupled with pervasive mobility and the proliferation of IoT applications, make the trustworthiness of collected data and the users' privacy the most important requirements to the success of these applications. Data trustworthiness is mandatory to build a chain of trust on a decision process taken according to IoT data and edge computations. This is very similar to the Web 2.0 scenario, where the plague of fake news and fake data substantially decreased the quality of decision processes, such as, for instance, in case of Twitter bots able to support or defame a specific product with high rates of success.

Traditional solutions to data validation mostly focused on assurance or reputation techniques. On one side, the system generating data undergoes an assurance process based on testing or monitoring; on the other side, the reputation of the entity owning the system is evaluated. Both approaches are however not viable in complex cyber–physical scenarios based on smart devices for different reasons: *i)* devices are usually resource constrained and cannot therefore be the target of extensive testing/monitoring; *ii)* IoT systems have fuzzy perimeters that are difficult to evaluate using traditional assurance techniques; *iii)* devices are often owned by unknown users, whose reputation cannot be easily evaluated; *iv)* people are often unintentional data providers, differently from Human-Provided Services where users consciously distribute services. Other approaches have provided solutions based on behavioral analysis, aimed to understand and differentiate human behaviors. In this scenario, there is an increasing need of privacy-aware solutions that evaluate the compliance of people to behavioral policies and, in turn, the trustworthiness of data collected through their devices.

In the first year of the project we focused on the need of a new human/user–centric approach in the context of smart devices. In the second year we extended the notion of behavioral assurance to cope with other user behaviors in IoT context, which is the most exposed environment in terms of risks of misbehaviors. This work is

currently under development starting from the novel assurance methodologies developed in the context of T1.4, which can be adapted for the user-centric domain. In addition to trustworthiness of data collected, privacy protection is unavoidable and must be carefully considered. We are working within Task 1.5 with University of Insubria to integrate our approach with solutions for users' privacy protection. The idea is to provide new privacy-aware security approaches (e.g., access control, identity management, authorization mechanisms) with verifiable behavior. We will continue working on this topic considering assurance techniques, blockchain, and the like, mainly in the context of Pilot "Security of Unmanned Aerial Systems (UAS)" to the aim of providing a privacy-aware location-based access control for UAV, and Pilot "e–Health Sector: Privacy and Data Protection" for managing emergency scenarios in smart homes.

8.4.11 Towards Trustworthy Remote Postal and Remote Electronic Voting systems

Contact: Christian Killer, Burkhard Stiller (UZH)

Trust in electoral processes is fundamental for democracies. Also, the management of citizen identities is crucial, because final tallies cannot be guaranteed without the assurance that an eligible voter cast every final vote. Therefore, two novel approaches were developed by UZH in the context of T1.5 of Concordia: (a) Proverum and (b) Provotum.

First, (a) Proverum focueses on establishing a basis for a hybrid public verifiability of voting, [79], an approach combining a private environment based on private permissioned Distributed Ledgers with a public environment based on public Blockchains (BC), and (2) describes the application of the Proverum architecture to the Swiss Remote Postal Voting system, mitigating threats present in the current system [78], and (3) addresses successfully the decentralized identity management in a federalistic state. The next by UZH will include the design of a voting scheme that introduces verifiability to the Swiss RPV system.

While the existence of Public Bulletin Boards (PBB) is often formulated as an assumption in related work on Remote Electronic Voting (REV) systems, Provotum [77] focuses on the practical design and architecture of such a PBB, including its distributed execution. Further, Provotum leverages a public permissioned BC as a PBB, where only authorized entities can sign blocks, while the general public can verify all BC data. Therefore, Provotum defines a new and fully decentralized BC-based REV system, which deploys a permissioned BC as a PBB and allows for the explicit distribution of trust across different permissioned BC nodes. Provotum is operated in a fully distributed fashion by using Smart Contracts (SC), Distributed Key Generation (DKG), Homomorphic Encryption (HE), and Cooperative Decryption (CD), as well as employing client-side encryption, which enables ballot secrecy, while the BC forms an audit trail, enabling public and End-to-end

Verifiability (E2E-V). [77]. The next steps in UZH will perform a detailed security analysis and implement Mixnets as a cryptographic basis and introduce Receipt-Freeness to avoid vote selling.

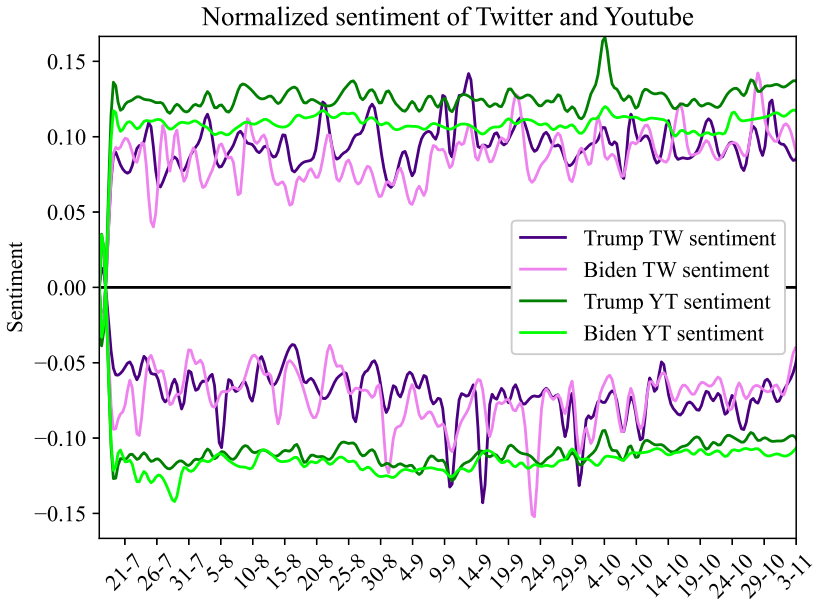

Figure 22: Normalized sentiment of Twitter and YouTube

8.4.12 Analysis of Twitter and YouTube during US elections 2020

Contact: Despoina Antonakaki (FORTH), Sotiris Ioannidis (FORTH)

The upcoming November 2020 presidential elections in the United States have caused extensive discussions on social media. A part of the content on US elections is organic, coming from users discussing their opinions of the candidates, political positions, or relevant content presented on television. Another significant part of the content generated originates from organized campaigns, both official and by astroturfing.

Initially, we obtain approximately 17.5 M tweets containing 3M users, based on prevalent hashtags related to US election 2020, as well as the related YouTube links, contained in the Twitter dataset, likes, dislikes and comments of the videos and conduct volume, sentiment and graph analysis on the communities formed. Particularly, we study the daily traffic per prevalent hashtags and show the evolution of the retweet graph from July to September 2020, highlighting the two main entities ('Biden' and 'Trump') contained in our dataset. Additionally, we gather the related YouTube links contained in the previous dataset and perform sentiment analysis. The results on sentiment analysis on the Twitter corpus and

the YouTube metadata gathered, show the positive and negative sentiment for the two entities throughout this period. The results of sentiment analysis indicate that 45.7% express positive sentiment towards Trump in Twitter and 33.8% positive sentiment towards Biden, while 14.55% of users express positive sentiment in YouTube metadata gathered towards Trump and 8.7% positive sentiment towards Biden. In Fig. 22 we the normalized sentiment of Twitter and YouTube for both entities. Also, in Fig. 23 we show the retweet graph for 3/11/2020, with red the trump related retweets and in blue the Biden related retweets.

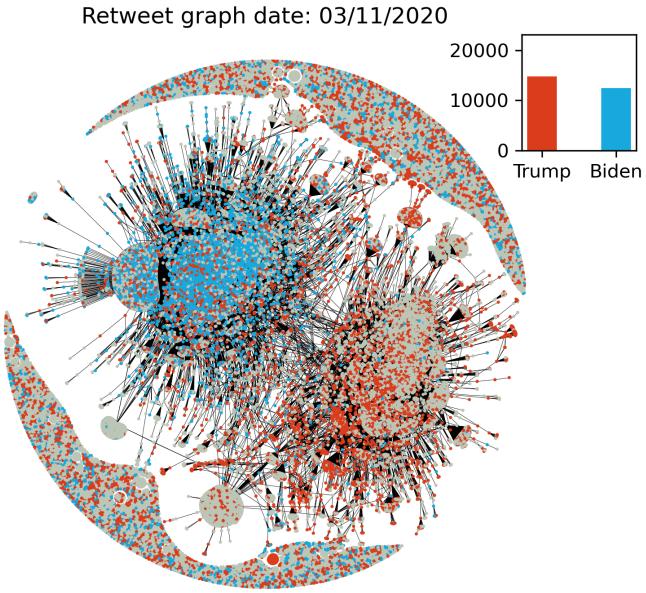

Figure 23: The retweet graph for 3/11/2020

Our analysis fills the gap between the connection of offline events and their consequences in social media by monitoring important events in real world and measuring public volume and sentiment before and after the event in social media [132]. This works applies to the data and user-centric security in CONCORDIA.

8.4.13 The impact of State–sponsored Trolls during the 2016 US Presidential Election Discourse

Contact: Nikos Salamanos, Michael Sirivianos (CUT)

In [124] we have revised and extended the study related to the impact of troll activities on the virality of the ambiguous political information that had been shared on Twitter during the 2016 US Presidential election (https://arxiv.org/abs/1910.00531). For that purpose, we follow a more general approach studying all the retweet cascades. Specifically: (i) We extend the analysis of the diffusion cascades by recon-

structing the series of retweets from our initial Twitter dataset, where the authentic retweet-labels are missing from the data. (ii) We apply graph structural analysis in order to qualify the position of troll accounts/nodes on the macroscopic level of interactions. (iii) We construct the *flow graphs*, which represent all possible paths of influence between the users that are present in the graph. iv) We use Shapley-Value based centrality measure in order to quantify the users' influence on the cascade-graphs. v) We identify the time-inferred diffusion cascade trees and we measure their structural virality as well as the impact of each user on them. vi) Finally, we perform top–k analysis in order to identify the influential users based on all viral cascades. The position of troll accounts on the top–k is a strong indicator of the influence they had on the diffusion of information. We provide a global influence ranking of all Twitter accounts and we find that one troll account appears in the top–100 and four in the top–1000. This along with other findings presents evidence that the driving force of virality and influence in the network came from regular users – users who have not been classified as trolls by Twitter. On the other hand, we find that on average, troll accounts were tens of times more influential than regular users were. Moreover, 23% and 22% of regular accounts in the top–100 and top–1000 respectively, have now been suspended by Twitter. This raises questions about their authenticity and practices overall, during the 2016 US presidential election.

This work is related to the third pillar of T1.5.

8.4.14 Discovery and classification of Twitter bots

Contact: Alexander Shevtsov (FORTH), Sotiris Ioannidis (FORTH)

A very large number of people use Online Social Networks daily. Such platforms thus become attractive targets for agents that seek to gain access to the attention of large audiences, and influence perceptions or opinions. Botnets, collections of automated accounts controlled by a single agent, are a common mechanism for exerting maximum influence. Botnets may be used to better infiltrate the social graph over time and to create an illusion of community behavior, amplifying their message and increasing persuasion.

Our efforts have been focused on the on the investigation of Twitter botnets, their behavior, their interaction with user communities and their evolution over time. We analyzed a dense crawl of a subset of Twitter traffic, amounting to nearly all interactions by Greek-speaking Twitter users for a period of 36 months. We detected over a million events where seemingly unrelated accounts tweeted nearly identical content at nearly the same time. We filtered these concurrent content injection events and detected a set of 1,850 accounts that repeatedly exhibit this pattern of behavior, suggesting that they are fully or in part controlled and orchestrated by the same software. In Fig. 24 we plot the bot graph of 2020.

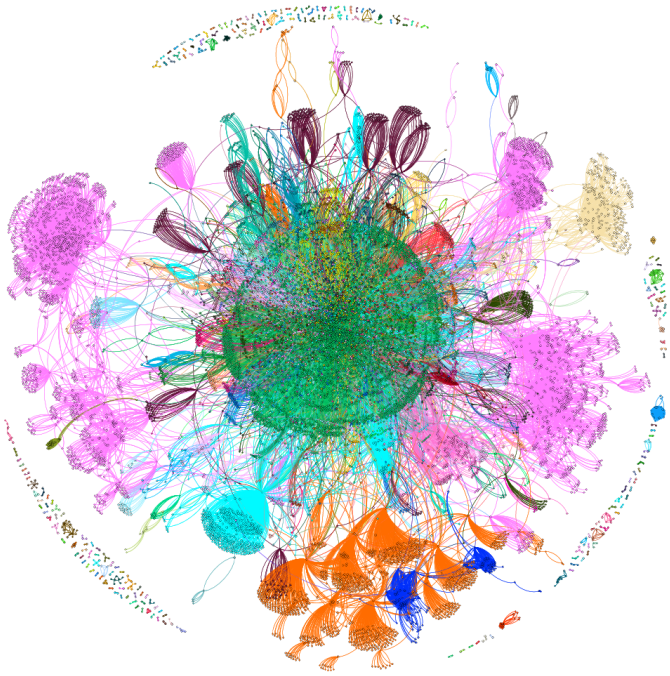

Figure 24: The bot graph for 2020

We found botnets that appear for brief intervals and disappear, as well as botnets that evolve and grow, spanning the duration of our dataset. We analyze statistical differences between bot accounts and human users, as well as botnet interaction with user communities and Twitter trending topics [133]. This works applies to the data and user-centric security in CONCORDIA.

8.4.15 Streaming Machine Learning Framework for Online Aggression Detection

Contact: Herodotos Herodotou (CUT), Despoina Chatzakou (ITI-CERTH), Nicolas Kourtellis (Telefonica)

The rise of online aggression on social media is evolving into a major point of concern. Several machine and deep learning approaches have been proposed recently for detecting various types of aggressive behavior. However, social media are fast paced, generating an increasing amount of content, while aggressive behavior evolves over time. In this work, we introduce the first, practical, real-time framework for detecting aggression on Twitter via embracing the streaming machine learning paradigm. Our method adapts its ML classifiers in an incremental fashion as it receives new annotated examples and is able to achieve the same (or even higher) performance as batch- based ML models, with over 90% accuracy,

precision, and recall. At the same time, our experimental analysis on real Twitter data reveals how our framework can easily scale to accommodate the entire Twitter Firehose (of 778 million tweets per day) with only 3 commodity machines. Finally, we show that our framework is general enough to detect other related behaviors such as sarcasm, racism, and sexism in real time.

We present this work as a short, preliminary investigation in IEEE ICDE [65], and a longer version in IEEE Big Data paper [64].

This work is generally related to the third pillar of T1.5.

8.4.16 Characterizing and detecting inappropriate videos targeting young children

Contact: Kostantinos Papadamou (CUT), Michalis Sirivianos (CUT), Nicolas Kourtellis (Telefonica)

A large number of the most-subscribed YouTube channels target children of very young age. Hundreds of toddler-oriented channels on YouTube feature inoffensive, well produced, and educational videos. Unfortunately, inappropriate content that targets this demographic is also common. YouTube's algorithmic recommendation system regrettably suggests inappropriate content because some of it mimics or is derived from otherwise appropriate content. Considering the risk for early childhood development, and an increasing trend in toddler's consumption of YouTube media, this is a worrisome problem. In this work, we build a classifier able to discern inappropriate content that targets toddlers on YouTube with 84.3% accuracy, and leverage it to perform a large-scale, quantitative characterization that reveals some of the risks of YouTube media consumption by young children. Our analysis reveals that YouTube is still plagued by such disturbing videos and its currently deployed counter-measures are ineffective in terms of detecting them in a timely manner. Alarmingly, using our classifier we show that young children are not only able, but likely to encounter disturbing videos when they randomly browse the platform starting from benign videos.

We presented this work in AAAI ICWSM [106], in which it received Honorable mention.

This work is generally related to the third pillar of T1.5.

8.4.17 Understanding the Incel Community on YouTube

Contact: Kostantinos Papadamou, Nikos Salamanos, Michael Sirivianos (CUT)

In [108], we studied the Incel community on YouTube. We focused on the evolution of this community over the last decade as well as on understanding whether the YouTube recommendation algorithm drives users towards Incel-related videos.

For that purpose, we collected videos shared on Incel-related communities within Reddit, in order to perform data-driven characterization of the material posted on YouTube. Among other things, we find that the Incel community on YouTube is getting traction and that during the last decade the number of Incel-related videos and comments rose substantially. Moreover, we quantify the probability that a user will encounter an Incel–related video by virtue of YouTube recommendation algorithm. Within five hops when starting from a non-Incel-related video, this probability is 1 in 5, which is alarmingly high as such content is likely to share toxic and misogynistic views.

This work is related to the third pillar of T1.5.

8.4.18 Analysis of YouTube's Pseudoscientific Video Recommendations

Contact: Kostantinos Papadamou, Nikos Salamanos, Michael Sirivianos (CUT)

YouTube has revolutionized the way people discover and consume videos, becoming one of the primary news sources for Internet users. Since content on YouTube is generated by its users, the platform is particularly vulnerable to misinformative and conspiratorial videos. Even worse, the role played by YouTube's recommendation algorithm in unwittingly promoting questionable content is not well understood, and could potentially make the problem even worse. This can have dire real-world consequences, especially when pseudoscientific content is promoted to users at critical times, e.g., during the COVID-19 pandemic. In [107], we set out to characterize and detect pseudoscientific misinformation on YouTube. We collect 6.6K videos related to COVID-19, the flat earth theory, the anti-vaccination, and anti-mask movements; using crowd–sourcing, we annotate them as pseudoscience, legitimate science, or irrelevant. We then train a deep learning classifier to detect pseudoscientific videos with an accuracy of 76.1%. Next, we quantify user exposure to this content on various parts of the platform (i.e., a user's homepage, recommended videos while watching a specific video, or search results) and how this exposure changes based on the user's watch history. We find that YouTube's recommendation algorithm is more aggressive in suggesting pseudoscientific content when users are searching for specific topics, while these recommendations are less common on a user's homepage or when actively watching pseudoscientific videos. Finally, we shed light on how a user's watch history substantially affects the type of recommended videos.

This work is related to the third pillar of T1.5.

8.4.19 Detecting disinformation diffusion in Twitter: COVID–19 data

Contact: Nikos Salamanos, Michael Sirivianos (CUT)

In the context of the third pillar of T1.5, CUT is developing a machine learning method for detecting disinformation diffusion in Twitter. As a first step, we collect the tweets IDs from github repository (https://github.com/echen102/COVID-19-TweetIDs). The repository contains an ongoing collection of tweets IDS associated with the novel coronavirus COVID-19, which commenced on January 28, 2020. The tweets IDs are collected from specified accounts and also collect in real-time tweets that mention specific keywords. The second step starts at the end of each month and includes the collection of all the extended information of all tweets IDs of the previous month. During the third step, we analyze all of the newly collected tweets information, extract any included URLs, identify those that are using URL shortening services and expand them into their long URL format. At this step we also collect all the keywords associated with the detected URLs. Overall the dataset includes around 850 Million tweets.

9 Organization of the Scientific Community and Events

The Description of Action identifies four (SMART) objectives for WP1 (see Section 1). Two of these objectives concern the organization of the scientific community and events:

- Organization of scientific events in the area of cybersecurity, including a dedicated annual European cybersecurity conference;
- Leading role in the organization of the scientific community, outreach to different target audiences, including public media and the general public.

The following sections will outline progress and achievements connected to these objectives. An observation that carries over from the first year is that CONCORDIA activities appear to focus on venues where security is *applied* within a specific context. The ACM Internet Measurement Conference (IMC) is a good example of such a venue. IMC publishes excellent measurement-based research relating to (the security of) networked services and the Internet.

The focus on *applied security* provides a strong indication that real-world security problems and challenges inspire CONCORDIA research and moreover that CONCORDIA activities aim to have an impact in this area.

During the first year of the project, work was set in motion to co-organize two PhD schools: the 4th International NeCS PhD Winter School and the IFIP Summer School on Privacy and Identity Management. (We previously reported on these efforts in deliverable D1.1.) The PhD Winter School took place in Trento, Italy, in January, 2020. The Summer School on Privacy and Identity Management was rescheduled to September 2020 as a fully virtual event due to the ongoing pandemic. These schools were not only co-organized by personnel from CONCORDIA partners (MUNI, UT, UZH) – other CONCORDIA researchers were also involved in creating the content taught at these schools. For example, Mattijs Jonker and Raffaele Sommese (UT) delivered a tutorial on "Internet Security Research using Public Measurement Data and Apache Spark." Jeroen van der Ham (UT) delivered a class on "Cyber Security & Incident Response."

Conference Tutorials & Hands-on Workshops

Efforts to share practical experience and skills was not limited to the PhD schools mentioned above. CONCORDIA researchers Mattijs Joner (UT) and Nils Rodday (CODE) chaired a hands-on tutorial on "Reliable measurements with BGP and RPKI" at the 2020 IEEE/IFIP Network Operations and Management Symposium 2020.[40]

[40] https://noms2020.ieee-noms.org/program/tutorials

9.1 Organization of Scientific Events

CONCORDIA members continue to be very active in the organization of scientific conferences. These conferences include (in alphabetic order):

- 10th ACM Conference on Data and Application Security and Privacy
- 11th ACM Conference on Data and Application Security and Privacy
- 12th IEEE International Conference on Cloud Computing
- 12th IEEE International Conference on Web Services
- 16th International Conference on emerging Networking EXperiments and Technologies
- 35th International Conference on ICT Systems Security and Privacy Protection
- 36th IEEE International Conference on Data Engineering
- 6th IEEE/IFIP Workshop on Security for Emerging Distributed Network Technologies
- ACM/IRTF Applied Networking Research Workshop
- IEEE 19th International Conference on Industrial Informatics - Track 3. Safety and Security in Industrial Applications
- IEEE Services Doctoral Symposium 2020
- Passive and Active Measurement conference 2020
- The 13th International Symposium on Foundations and Practice of Security
- The Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications

Membership of Technical Program Committees (TPC)

Appendix C contains a table with all TPC related efforts of CONCORDIA researchers during the year 2020. As can be seen, researchers have been active as TPC member in well over 70 different conferences.

9.2 Organization of Scientific Community

The organization of the scientific community involves the following three activities:

- Chairing professional organizations (ACM, IEEE and IFIP);
- Editing scientific journals;

- Membership of steering committee.

Chairing Professional Organizations

The scientific community is organized by professional organizations such as ACM, IEEE, and IFIP. Various CONCORDIA researchers hold positions within these organizations. The following positions are (or have continued to be through the second year) held by CONCORDIA researchers:

- Chair of IFIP TC6[41] (Burkhard Stiller, UZH)
  The Technical Committee 6 (TC6), *Communications Systems*, is one of the largest TCs within IFIP in terms of activities and revenues. TC6 has nine Working Groups (WGs), as well as a number of Special Interest Groups (SIGs), the majority of which are concerned either with specific aspects of communications systems themselves or with the application of communications systems. In addition, one WG focuses on communications in developing countries. TC6 meets twice a year, in spring and fall.
- Chair of IFIP TC6 Working Group 6.6[42] (Rémi Badonnel, UL)
  The Working Group 6 of IFIP TC6 focuses on the management of network and distributed systems. Management is defined in five functional areas – the so-called "FCAPS" areas. These involve: Fault management, Configuration management, Accounting management, Performance management and Security management. Security management has become the greatest challenge.

Editing Scientific Journals

Nine members, i.e., researchers, of the CONCORDIA consortium acted as journal editors in 2020, with varying roles including but not limited to as special issue editor, special chief editor, and associate editor. They were active for the following 25 journals:

- ACM DTRAP – Special Issue on Vulnerabilities
- ACM Transactions on Data Science
- ACM Transactions on Privacy and Security
- Communications Magazine
- Computer Communications Review
- Cybersecurity and Privacy of Frontiers in Big Data
- Data Science and Engineering

[41] https://ifip.informatik.uni-hamburg.de/ifip/tc/6
[42] https://ifip.informatik.uni-hamburg.de/ifip/tc/6/wg/6.6/officers

- Digital Threats: Research and Practice
- Digital Trust: Trust Management in the Cyberspace, IEEE Internet computing
- IEEE Internet Computing
- IEEE Transactions on Big Data
- IEEE Transactions on Cloud Computing
- IEEE Transactions on Emerging Topics in Computing
- IEEE Transactions on Information Forensics and Security
- IEEE Transactions on Network and Service Management
- IEEE Transactions on Service Computing
- Information Security Journal: A Global Perspective
- International Journal of Cooperative Information Systems
- International Journal of Network Management
- Journal of Network and Systems Management
- Secure, Efficient Cyber-Physical Systems and Wireless Sensors
- Sensors
- Special issue on Computing for Autonomy: Latency, Power, Resilience
- Sustainability
- Synthesis Lectures on Security, Privacy and Trust

Steering Committee Membership

CONCORDIA researchers are member of the Steering Committee for the following conferences:

- IFIP TMA – Traffic Measurement and Analysis Conference[43] (Anna Sperotto, UT)
- CAIDA WOMBIR – Workshop on Overcoming Measurement Barriers to Internet Research[44] (Mattijs Jonker, UT)

[43] https://tma.ifip.org/
[44] https://www.caida.org/workshops/wombir/

Non-Academic Conferences

On occasion, CONCORDIA researchers are also involved in the organization of non-academic events. We list such events here, provided the events have a strong focus on cybersecurity, and aim for knowledge sharing and technological advancement.

- FIRST 2020 – The Forum of Incident Response and Security Teams[45] (Jeroen van der Ham, UT)
- MCH 2021 – May Contain Hackers[46] (Jeroen van der Ham, UT)

[45] https://www.first.org/conference/2020/
[46] https://mch2021.org/

10 Contributions to Standards and Open Research Data

One of the objectives of WP1 (see Section 1) is to contribute to standardization, open research data and code, shared via systems such as GitHub. This section outlines the achievements in this area by WP1 researchers.

It should be noted that CONCORDIA has special tasks for standardization as well as open data:

- Task 5.3: Certification and Standardization activities;
- Task 6.4: Data management.

A more complete overview of all CONCORDIA activities related to standardization and open data is therefore included in deliverables of WP5 and WP6.

10.1 Standardization Activities Performed by WP1 Researchers

Standardization efforts takes multiple years from the initial research until the establishment of a full standard. The activities identified in this subsection are therefore activities that started years before, but obtained important outcome in 2020 and have some relationship with current WP1 research.

Recommendations for DNS Privacy Service Operators (RFC 8932 / BCP 232)

RFC 8932 [42] defines a comprehensive set of best practices for guarding user privacy for DNS resolver operators. Privacy of DNS resolution by clients has had the attention of the IETF since the so-called Snowden revelations. RFC 7626 contains an analysis of privacy risks in the DNS and the DPRIVE working group was formed to address the challenge of reducing these risks. Several RFCs published by this working group introduce encrypted transport for client-to-resolver traffic (chiefly RFC 7858 [73], which introduces DNS transport over TLS, colloquially referred to as DNS-over-TLS or DoT) and specify additional extensions to, for example, further protect the traffic against analysis by adding padding to hide payload sizes (i.e. RFC 7830 [88] that introduces EDNS0 padding and RFC 8467 [89] that provides guidance on how to use padding). In addition to this, a more recent standard from a dedicated working group introduced DNS-over-HTTPS (or DoH) transport for client-to-resolver DNS queries.

In RFC 8932, we provide comprehensive guidance on how to operate a DNS resolver service that implements privacy-conscious transports. The RFC not only highlights how to choose protocol settings for optimal privacy, but also discusses resolver configuration choices and specifies a framework for a Recursive Operator Privacy statement, which operators can use to outline their privacy policy.

Note: Roland van Rijswijk-Deij is listed in this RFC with his NLnet Labs affiliation, but he also worked on this standard from his position at the UT.

Recommendations for Secure Use of Transport Layer Security (RFC 7525bis / BCP 195)

Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS) are widely used to protect data exchanged over application protocols such as HTTP, SMTP, and SIP. Over the last years, several serious attacks on TLS have emerged, including attacks on commonly used cipher suites and their modes of operation. RFC 7525 [130] (2015) provides recommendations for improving the security of deployed services that use TLS and DTLS. These recommendations were formalized during the transition to TLS v1.2.

Nowadays, the transition is largely complete, and TLS v1.3 has become widely available. Under these new circumstances, we believe new and updated guidance is needed, which we propose in RFC 7525bis [131] as part of our standarization efforts towards securing networked applications.

10.2 Open Research Data Provided by WP1 Researchers

IoT-deNAT: Outbound flow-based network traffic data of IoT and non-IoT devices behind a home NAT

A Zenodo hosted repository[47], comprised of NetFlow records which capture the outbound network traffic of 8 commercial IoT devices and 5 non-IoT devices, collected during a period of 37 days in a lab at Ben-Gurion University of The Negev. The dataset was collected in order to develop a method for telecommunication providers to detect vulnerable IoT models behind home NATs. Each NetFlow record is labeled with the device model which produced it; for research reproducibilty, each NetFlow is also allocated to either the ”training” or ”test” set, in accordance with the partitioning described in [91].

Responses to AHP-style questionnaires regarding the publication “D-Score: A Novel Expert-Based Method for Assessing the Detectability of IoT-Related Cyber-Attacks”

In order to assess in advance the detectability of IoT-related cyber-attacks by anomaly-based network intrusion detection systems, at Ben-Gurion University of the Negev we developed an expert-based method which relies on the AHP methodology. The online AHP-style questionnaire can be found at: https://bguprivacysurvey.limequery.com/537457. The dataset that we made public via Zenodo includes the (anonymous) responses of 40 cyber-security researchers and practitioners, covering 4 IoT attack scenarios. The related manuscript [90] has been submitted to a leading journal for review.

[47] https://doi.org/10.5281/zenodo.3924770

Tracking the deployment of TLS 1.3 on the Web

As part of efforts to track the deployment, uptake and use of Transport Layer Security v1.3, we recently combined and analyzed data from active domain scans, passive monitoring of large networks, and a crowd-sourcing effort on Android devices [71]. We released the code of the scanning software, scripts for data analysis, and measurement data, where possible. These resources can be found at https://tls13.globalsecuritylabs.org/.

User Exposure to Illicit Cryptocurrency Mining on the Web

A few years ago, a new malware phenomenon appeared on the Internet: browser-based cryptocurrency mining. This attack vector allows miscreants to exploit the CPU time of unsuspecting website visitors. We investigated how many users were exposed to such sites and whether there was a real risk for users given common browsing behavior [72].

We use passive measurements to obtain an accurate picture of user exposure. Some of our passive measurements rely on input from active scans and classifications from blocklists to identify mining sites. To support open science, we have made the software and non-privacy sensitive data available at https://github.com/retrocryptomining/data.

10.3 Research Tools & Software Provided by WP1 Researchers

LoRadar: LoRa sensor network monitoring through passive packet sniffing

LoRa (Long Range) is a technology for Low Power WAN technologies (LPWANs) that have low power consumption and support longer transmission ranges. LoRa has recently gained significant popularity due to its ease of deployment and it is used to facilitate data communication in IoT deployments. It is important to develop measurement tools that allow investigation into the performance, utilization, and security aspects of LoRa networks.

In a recent paper [27], we present *LoRadar*, a passive packet sniffing framework for LoRa's Medium Access Control (MAC) protocol, LoRaWAN. LoRadar can provide key insight into LoRa deployments. LoRadar was released to the public (see: https://github.com/loradar/loradar_tool).

11 Conclusions and Outlook

The second year of the CONCORDIA project has been a good one for WP1. In terms of the main objective, stimulating the publication of scientific results in the broad field of cybersecurity, WP1 researchers did exceptionally well. All of the five WP1 tasks saw the successful execution of research directives set in year one, as well as the continuation of ongoing efforts. As a result, over 100 additional papers were published during the second year, many of them in renowned conferences and journals. This outclasses a first year that was successful in its own right (50 papers published) and is gives a promising outlook for the next two years in terms of writing high-quality papers.

The dissemination of successful WP1 research went beyond academic conferences and journals alone. CONCORDIA researchers shared their findings at a number of non-academic events. Such events are typically attended by industry and the wider Internet community (network operators, Internet engineers and architecture designers). As such, targeting this crowd allows WP1 output to stimulate improvements on, e.g., an operational level. A concrete example of this is task T1.2's investigation into DNS zone administration errors, which following dissemination at DNS-OARC and RIPE80, among others, prompted changes by DNS operators to improve DNS security and stability. Other examples of dissemination to a wider audience are the white papers by T1.5 on the impact of Twitter and YouTube on the US elections, and the blog on 'Ethics, Responsibilities and Vulnerabilities' at FIRST

During the first year of the project, collaboration between WP1 tasks and the pilots was explored. Possible synergetic research activities were chartered. Over the second year, some of the planned work came to fruition. A concrete example is the joint work between network-centric security research (T1.2) and the DDoS Clearing House for Europe pilot (T3.2). Actionable intelligence from T1.2 can now feed into the pilot's architecture to enrich DDoS attack signatures. And further improvements are planned for year three. Another example is the strong collaboration that has been established between application-centric security research (T1.4) and the Threat Intelligence for the Telco Sector pilot (T2.1), where T1.4's efforts in investigating application security has been instrumental in making sure that requirements set by T2.1 are met.

When it comes to the organization of scientific events, CONCORDIA researchers also did well. Researchers sat on the technical program committees of over 70 different conferences during the second year. Moreover, they were involved in the organization of 15 academic events and even a handful non-academic yet cybersecurity related events. Virtually all of these events had to be held virtually, which created challenges.

In 2020 new activities led to a first draft of a roadmap for cyber security research. The starting point of this roadmap is the societal challenge that Europe is facing,

which is our digital sovereignty. Derived from that are a list of research challenges to improve the security of devices, networks, systems, applications and users. Although some of these challenges have already been identified before at the European level (such as microcontrollers, cloud services and quantum), it is remarkable that until now limited attention has been paid on our core infrastructure, which is the Internet. One of the main conclusions of the CONCORDIA roadmap is therefore that additional research is needed to design a 'Responsible Internet', a novel security-by-design extension of the Internet that offers users better grip on dependencies, thus increasing trust. It should be noted that the notion of a Responsible Internet goes much further than 5G security, which focusses primarily on the cellular access part of the Internet. The complete CONCORDIA roadmap can be found in Deliverable 4.4, but the research specific parts of the roadmap have also been included in this deliverable as part of the individual task descriptions.

The second year of the project coincided with the outbreak of the COVID-19 pandemic. From an academic perspective alone, the pandemic has brought about a number of challenges. When it comes to the impact on WP1 activities, we conclude that the pandemic's effect on the execution of research has been limited but certainly not nil. Many of WP1's research activities do not involve labs, but on the rare occasion that a lab had to be shut down, research was negatively affected. From a mental wellbeing side, there are more unknowns. Junior researchers often live abroad and have a smaller social safetynet. Many of them have been isolated due to university closure. We suspect that some effects on their well-being of are likely to be serious. The pandemic has also made attending events in-person much harder, which hinders opportunities to meet peers and develop a professional network.

For 2021 WP1 has a number of goals. First, we would like to strengthen the support for young PhDs, and help them to shape their careers. Whenever possible, we would like to collaborate in this area with the three other pilot projects, as well as other organizations, to ensure that support continues even after the end of this project. Second we would like to extend the collaboration with industry partners and increase the impact of our research. For that additional focus may be needed, for example in the areas of Threat Intelligence data, DDoS protection or more general the resilience and trustworthiness of our Internet. Third we would like to further investigate the role and behaviour of social media, to increase the security and privacy of our end users.

References

[1] CYST: Cybersecurity Discrete-Event Simulation Engine. https://muni.cz/go/565e43. Online; Accessed: 09-12-2020.

[2] Getting around non-executable stack (and fix). https://seclists.org/bugtraq/1997/Aug/63. Online; Accessed: 24-06-2020.

[3] Trace-Share: An open platform for creation and sharing of network traces. https://github.com/CSIRT-MU/Trace-Share. Online; Accessed: 09-12-2020.

[4] D. S. Aanchal Malhotra, Austin King and M. Zochowski. Payid protocol. Technical report, Ripple, June 2020.

[5] P. Agarwal, S. Joglekar, P. Papadopoulos, N. Sastry, and N. Kourtellis. Stop tracking me Bro! Differential Tracking of User Demographics on Hyper-Partisan Websites. In *Proceedings of The Web Conference 2020*, April 2020.

[6] G. Akiwate, M. Jonker, R. Sommese, I. Foster, G. M. Voelker, S. Savage, and K. Claffy. Unresolved issues: Prevalence, persistence, and perils of lame delegations. In *Proceedings of the ACM Internet Measurement Conference*, pages 281–294. ACM, 2020.

[7] C. Alexakos, C. Katsini, K. Votis, A. Lalas, D. Tzovaras, and S. D. Enabling digital forensics readiness for internet of vehicles. https://ewgt2020.eu/wp-content/uploads/extended_abstract/EWGT2020_abstract_215.pdf, 09 2020.

[8] Z. Alom, B. Carminati, and E. Ferrari. A deep learning model for twitter spam detection. *Online Social Networks and Media*, 18:100079, 2020.

[9] T. Alves. Trustzone: Integrated hardware and software security. *White paper*, 2004.

[10] M. Anisetti, C. Ardagna, N. Bena, and E. Damiani. Stay thrifty, stay secure: A vpn-based assurance framework for hybrid systems. In *International Conference on Security and Cryptography*, pages 98–109. ScitePress, 2020.

[11] M. Anisetti, C. Ardagna, B. Carminati, E. Ferrari, and C. Perner. Requirements and challenges for secure and trustworthy uas collaboration. In *TPS 2020*, 2020 (to appear).

[12] M. Anisetti, C. Ardagna, B. Carminati, C. Rondanini, E. Ferrari, and E. Damiani. A blockchain-based trustworthy certification process for composite services. In *IEEE International Conference on Services Computing (SCC)*, 2020 (to appear).

[13] M. Anisetti, C. Ardagna, E. Damiani, and P. Panero. A methodology for non-functional property evaluation of machine learning models. In *ACM Management of Emergent Digital EcoSystems (MEDES)*, 2020 (to appear).

[14] C. Ardagna, R. Asal, E. Damiani, N. El Ioini, and C. Pahl. Trustworthy IoT: An evidence collection approach based on smart contracts. In *Proc. of the 15th IEEE International Conference on Services Computing (SCC 2019)*, Milan, Italy, July 2019.

[15] T. Arnold, E. Gürmeriçliler, G. Essig, A. Gupta, M. Calder, V. Giotsas, and E. Katz-Bassett. (how much) does a private wan improve cloud performance? In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pages 79–88. IEEE, 2020.

[16] S. Bardin, T. Benoit, and J.-Y. Marion. Compiler and optimization level recognition using graph neural networks. In *Machine Learning for program analysis (MLPA workshop)*, Jan. 2021.

[17] L. M. Bertholdo, J. M. Ceron, L. Z. Granville, G. C. Moura, C. Hesselman, and R. van Rijswijk-Deij. Bgp anycast tuner: Intuitive route management for anycast services. 2020.

[18] A. Bierman, M. Björklund, and K. Watsen. RESTCONF Protocol. RFC 8040, YumaWorks, Tail-f Systems, Juniper Networks, Jan. 2017.

[19] M. Björklund. The YANG 1.1 Data Modeling Language. RFC 7950, Tail-f Systems, Aug. 2016.

[20] T. Bletsch, X. Jiang, V. W. Freeh, and Z. Liang. Jump-oriented programming: a new class of code-reuse attack. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pages 30–40, 2011.

[21] P.-O. Brissaud, J. François, I. Chrisment, T. Cholez, and O. Bettan. Transparent and Service-Agnostic Monitoring of Encrypted Web Traffic. *IEEE Transactions on Network and Service Management*, 16(3):842–856, Sept. 2019.

[22] P.-O. Brissaud, J. François, I. Chrisment, T. Cholez, and O. Bettan. Encrypted HTTP/2 Traffic Monitoring: Standing the Test of Time and Space. In *WIFS2020 - IEEE International Workshop on Information Forensics and Security*, New-York/Virtual, United States, Dec. 2020.

[23] B. Carminati, P. Colombo, E. Ferrari, and G. Sagirlar. Enhancing user control on personal data usage in internet of things ecosystems. In *2016 IEEE International Conference on Services Computing (SCC)*, pages 291–298, 2016.

[24] J. Casas-Roma, J. Salas, F. D. Malliaros, and M. Vazirgiannis. k-degree anonymity on directed networks. *Knowledge and Information Systems*, pages 1743–1768, 2019.

[25] N. Chalkiadakis, D. Deyannis, D. Karnikis, G. Vasiliadis, and S. Ioannidis. The million dollar handshake: Secure and attested communications in the cloud.

[26] B. Cheng, J. Ming, E. Leal, H. Zhang, J. Fu, G. Peng, and J.-Y. Marion. Extracting Executable Payloads From Packed Malware: Import Table Reconstruction via Hardware-Assisted API Micro Execution. In *USENIX Security '21 Fall*, Feb. 2021.

[27] K. N. Choi, H. Kolamunna, A. Uyanwatta, K. Thilakarathna, S. Seneviratne, R. Holz, M. Hassan, and A. Y. Zomaya. Loradar: Lora sensor network monitoring through passive packet sniffing. *SIGCOMM Computer Communication Review*, 50(4):10–24, 2020.

[28] D. Cohen, Y. Mirsky, M. Kamp, T. Martin, Y. Elovici, R. Puzis, and A. Shabtai. Dante: A framework for mining and monitoring darknet traffic. In *European Symposium on Research in Computer Security*, pages 88–109. Springer, 2020.

[29] P. Colombo and E. Ferrari. Access control enforcement within mqtt-based internet of things ecosystems. In *Proceedings of the 23nd ACM on Symposium on Access Control Models and Technologies*, SACMAT '18, page 223–234, New York, NY, USA, 2018. Association for Computing Machinery.

[30] P. Colombo and E. Ferrari. Evaluating the effects of access control policies within nosql systems. *Future Generation Computer Systems*, 114:491 – 505, 2021.

[31] P. Colombo, E. Ferrari, and S. Salvia. Mammoth: Monitoring the abac monitor of mqtt-based internet of things ecosystems. SACMAT '20, page 221–222, New York, NY, USA, 2020. Association for Computing Machinery.

[32] P. Colombo, E. Ferrari, and E. D. Tümer. Regulating data sharing across mqtt environments. *Journal of Network and Computer Applications*, 174:102907, 2021.

[33] E. Damiani and C. A. Ardagna. Certified machine-learning models. In *SOFSEM 2020: Theory and Practice of Computer Science*, pages 3–15, Cham, 2020. Springer International Publishing.

[34] E. Damiani and C. A. Ardagna. Certified machine-learning models. In *International Conference on Current Trends in Theory and Practice of Informatics*, pages 3–15. Springer, 2020.

[35] A. De Carli, M. Franco, A. Gassmann, C. Killer, B. Rodrigues, E. Scheid, D. Schoenbaechler, and B. Stiller. Wetrace–a privacy-preserving mobile covid-19 tracing approach and application. *arXiv preprint arXiv:2004.08812*, 2020.

[36] W. de Vries, R. de O. Schmidt, and A. Pras. Anycast and its potential for ddos mitigation. 06 2016.

[37] W. B. de Vries, S. Aljammāz, and R. van Rijswijk-Deij. Global-Scale Anycast Network Management with Verfploeter. In *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–9. IEEE, 2020.

[38] D. Deyannis, D. Karnikis, G. Vasiliadis, and S. Ioannidis. An enclave assisted snapshot-based kernel integrity monitor. In *Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking*, pages 19–24, 2020.

[39] D. Deyannis, E. Papadogiannaki, G. Kalivianakis, G. Vasiliadis, and S. Ioannidis. Trustav: Practical and privacy preserving malware analysis in the cloud. In *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy*, pages 39–48, 2020.

[40] M. Diamantaris, F. Marcantoni, S. Ioannidis, and J. Polakis. The seven deadly sins of the html5 webapi: A large-scale study on the risks of mobile sensor-based attacks. *ACM Transactions on Privacy and Security (TOPS)*, 23(4):1–31, 2020.

[41] M. Diamantaris, E. P. Papadopoulos, E. P. Markatos, S. Ioannidis, and J. Polakis. Reaper: Real-time app analysis for augmenting the android permission system. In *Proceedings of the Ninth ACM Conference on Data and Application Security and Privacy*, page 37–48, 2019.

[42] S. Dickinson, B. Overeinder, R. van Rijswijk-Deij, and A. Manik. Recommendations for DNS Privacy Service Operators. RFC 8932, Oct. 2020.

[43] C. Dietz, G. Dreo, A. Sperotto, and A. Pras. Towards adversarial resilience in proactive detection of botnet domain names by using mtd. In *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–5. IEEE, 2020.

[44] M. Drašar, S. Moskal, S. Yang, and P. Zat'ko. Session-level adversary intent-driven cyberattack simulator. In *2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*, pages 1–9, 2020.

[45] B. Dzogovic, B. Santos, N. Jacot, B. Feng, T. Van Do, et al. Secure healthcare: 5g-enabled network slicing for elderly care. In *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, pages 864–868. IEEE, 2020.

[46] R. Enns, M. Bjorklund, J. Schönwälder, and A. Bierman. Network Configuration Protocol (NETCONF). RFC 6241, Juniper Networks, Tail-f Systems, Jacobs University, Brocade, June 2011.

[47] G. Fidel, R. Bitton, and A. Shabtai. When explainability meets adversarial learning: Detecting adversarial examples using shap signatures. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020.

[48] A. P. Fournaris, C. Dimopoulos, and O. Koufopavlou. Profiling dilithium digital signature traces for correlation differential side channel attacks. In A. Orailoglu, M. Jung, and M. Reichenbach, editors, *Embedded Computer Systems: Architectures, Modeling, and Simulation*, pages 281–294, Cham, 2020. Springer International Publishing.

[49] A. P. Fournaris, C. Dimopoulos, K. Lampropoulos, and O. Koufopavlou. Anomaly detection trusted hardware sensors for critical infrastructure legacy devices. *Sensors (Switzerland)*, 20(11):3092, may 2020.

[50] A. P. Fournaris, C. Dimopoulos, K. Lampropoulos, and O. Koufopavlou. Anomaly detection trusted hardware sensors for critical infrastructure legacy devices. *Sensors*, 20(11):3092, 2020.

[51] A. P. Fournaris, A. Lalos, P. Kapsalas, and C. Koulamas. Decentralized, secure and cognitive architecture for automotive cyberphysical system of systems. In *2020 9th Mediterranean Conference on Embedded Computing (MECO)*, pages 1–5, 2020.

[52] A. P. Fournaris, A. Lalos, P. Kapsalas, and C. Koulamas. Decentralized, secure and cognitive architecture for automotive cyberphysical system of systems. In *2020 9th Mediterranean Conference on Embedded Computing (MECO)*, pages 1–5. IEEE, 2020.

[53] M. Franco, E. Sula, B. Rodrigues, E. Scheid, and B. Stiller. Protectddos: A platform for trustworthy offering and recommendation of protections. In *International Conference on the Economics of Grids, Clouds, Systems, and Services*, pages 28–40. Springer, 2020.

[54] M. F. Franco, B. Rodrigues, E. J. Scheid, A. Jacobs, C. Killer, L. Z. Granville, and B. Stiller. Secbot: a business-driven conversational agent for cybersecurity planning and management. In *2020 16th International Conference on Network and Service Management (CNSM)*, pages 1–7. IEEE, 2020.

[55] G. Gallopeni, B. Rodrigues, M. Franco, and B. Stiller. A practical analysis on mirai botnet traffic. In *2020 IFIP Networking Conference (Networking)*, pages 667–668. IEEE, 2020.

[56] A. Gehani and G. Kedem. RheoStat : Real-time Risk Management. In *Proc. of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID'2014)*, pages 15–17, 2004.

[57] G. Giakoumakis, E. Papadogiannaki, G. Vasiliadis, and S. Ioannidis. Pythia: Scheduling of concurrent network packet processing applications on heterogeneous devices. In *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pages 145–149. IEEE, 2020.

[58] V. Giotsas, I. Livadariu, and P. Gigis. A first look at the misuse and abuse of the ipv4 transfer market. In *International Conference on Passive and Active Network Measurement*, pages 88–103. Springer, 2020.

[59] V. Giotsas, G. Nomikos, V. Kotronis, P. Sermpezis, P. Gigis, L. Manassakis, C. Dietzel, S. Konstantaras, and X. Dimitropoulos. O peer, where art thou? uncovering remote peering interconnections at ixps. *IEEE/ACM Transactions on Networking*, 2020.

[60] M. Granderath and J. Schönwälder. A Resource Efficient Implementation of the RESTCONF Protocol for OpenWrt Systems. In *Proc. 17th IEEE/IFIP Network Operations and Management Symposium (NOMS 2020)*. IEEE, Apr. 2020.

[61] B. Green, R. Derbyshire, W. Knowles, J. Boorman, P. Ciholas, D. Prince, and D. Hutchison. {ICS} testbed tetris: Practical building blocks towards a cyber security resource. In *13th {USENIX} Workshop on Cyber Security Experimentation and Test ({CSET} 20)*, 2020.

[62] M. Hamad and V. Prevelakis. SAVTA: A Hybrid Vehicular Threat Model: Overview and Case Study. In *MDPI Information, Vol. 11*, 2020.

[63] M. Hamad, E. Regnath, J. Lauinger, V. Prevelakis, and S. Steinhorst. SPPS: Secure Policy-based Publish/Subscribe System for V2C Communication. In *Conference on Design, Automation and Test in Europe (DATE)*, 2021. Accepted.

[64] H. Herodotou, D. Chatzakou, and N. Kourtellis. A Streaming Machine Learning Framework for Online Aggression Detection on Twitter. In *IEEE International Conference of Big Data*, December 2020.

[65] H. Herodotou, D. Chatzakou, and N. Kourtellis. Catching them red-handed: Real-time Aggression Detection on Social Media. In *IEEE International Conference of Data Engineering (ICDE)*, April 2021.

[66] C. Hesselman, P. Grosso, R. Holz, F. Kuipers, J. H. Xue, M. Jonker, J. de Ruiter, A. Sperotto, R. van Rijswijk-Deij, G. C. Moura, et al. A responsible internet to increase trust in the digital world. *Journal of Network and Systems Management*, 28(4):882–922, 2020.

[67] C. Hesselman, M. Kaeo, L. Chapin, K. Claffy, M. Seiden, D. McPherson, D. Piscitello, A. McConachie, T. April, J. Latour, and R. Rasmussen. The dns in iot: Opportunities, risks, and challenges. *IEEE Internet Computing*, 24(4):23–32, 2020.

[68] A.-T. Hoang, B. Carminati, and E. Ferrari. Cluster-based anonymization of directed graphs. In *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*, pages 91–100, 2019.

[69] A.-T. Hoang, B. Carminati, and E. Ferrari. Cluster-based anonymization of knowledge graphs. In *18th International Conference on Applied Cryptography and Network Security*, 2020.

[70] A.-T. Hoang, B. Carminati, and E. Ferrari. Privacy-preserving sequentially publishing of knowledge graphs. In *37th IEEE International Conference on Data Engineering*, 2021.

[71] R. Holz, J. Hiller, J. Amann, A. Razaghpanah, T. Jost, N. Vallina-Rodriguez, and O. Hohlfeld. Tracking the deployment of tls 1.3 on the web: A story of experimentation and centralization. *SIGCOMM Compututer Communication Review*, 50(3):3–15, 2020.

[72] R. Holz, D. Perino, M. Varvello, J. Amann, A. Continella, N. Evans, I. Leontiadis, C. Natoli, and Q. Scheitle. A retrospective analysis of user exposure to (illicit) cryptocurrency mining on the web. In *Proc. of the 2020 Network Traffic Measurement and Analysis Conference (TMA)*, 2020.

[73] Z. Hu, L. Zhu, J. Heidemann, A. Mankin, D. Wessels, and P. Hoofman. Specification for DNS over Transport Layer Security (TLS). RFC 7858, May 2016.

[74] A. K. Iannillo and R. State. A proposal for security assessment of trustzone-m based software. In *2019 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. IEEE, 2019.

[75] M. Jonker, A. Sperotto, and A. Pras. Ddos mitigation: A measurement-based approach. In *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6. IEEE, 2020.

[76] C. Killer, B. Rodrigues, R. Matile, E. Scheid, and B. Stiller. Design and implementation of cast-as-intended verifiability for a blockchain-based voting system. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 286–293, 2020.

[77] C. Killer, B. Rodrigues, E. J. Scheid, M. Franco, M. Eck, N. Zaugg, A. Scheitlin, and B. Stiller. Provotum: A Blockchain-Based and End-to-End Verifiable Remote Electronic Voting System. In *IEEE 45th Conference on Local Computer Networks (LCN)*, Sidney, Australia, nov 2020. IEEE.

[78] C. Killer and B. Stiller. The Swiss Postal Voting Process and Its System and Security Analysis. pages 134–149. 2019.

[79] C. Killer, L. Thorbecke, B. Rodrigues, E. Scheid, M. Franco, and B. Stiller. Proverum: A hybrid public verifiability and decentralized identity management. *arXiv preprint arXiv:2008.09841*, 2020.

[80] J. Kohlrausch and E. A. Brin. Arima supplemented security metrics for quality assurance and situational awareness. *Digital Threats: Research and Practice*, 1(1), Mar. 2020.

[81] J. Kotak and Y. Elovici. Iot device identification using deep learning. In *13th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2020)*, pages 76–86, Cham, 2021. Springer International Publishing.

[82] N. Kourtellis, K. Katevas, and D. Perino. FLaaS: Federated Learning as a Service. In *ACM Distributed ML*, December 2020.

[83] A. S. Lalos, E. Vlachos, K. Berberidis, A. P. Fournaris, and C. Koulamas. Privacy preservation in industrial iot via fast adaptive correlation matrix completion. *IEEE Transactions on Industrial Informatics*, 16(12):7765–7773, 2020.

[84] M. Laštovička, S. Špaček, P. Velan, and P. Čeleda. Using TLS Fingerprints for OS Identification in Encrypted Traffic. In *2020 IEEE/IFIP Network Operations and Management Symposium (NOMS 2020)*, pages 1–6. IEEE Xplore Digital Library, 2020.

[85] E. Lear, R. Droms, and D. Romascanu. Manufacturer Usage Description Specification. RFC 8520, Cisco Systems, Google, Mar. 2019.

[86] H. Lee, A. Gireesh, R. van Rijswijk-Deij, T. Chung, et al. A Longitudinal and Comprehensive Study of the {DANE} Ecosystem in Email. In *29th {USENIX} Security Symposium ({USENIX} Security 20)*, pages 613–630. USENIX Association, 2020.

[87] L. Mauri, E. Damiani, and S. Cimato. Be your neighbor's miner: Building trust in ledger content via reciprocally useful work. In *2020 IEEE 13th International Conference on Cloud Computing (CLOUD)*, 2020 (to appear).

[88] A. Mayrhofer. The EDNS(0) Padding Option. RFC 7830, May 2016.

[89] A. Mayrhofer. Padding Policies for Extension Mechanisms for DNS (EDNS(0)). RFC 8467, Oct. 2018.

[90] Y. Meidan, D. Benatar, R. Bitton, and A. Shabtai. An expert-based method for assessing the detectability of iot-related cyber-attacks. *Manuscript submitted for publication*, 2020.

[91] Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai. A novel approach for detecting vulnerable iot devices connected behind a home nat. *Computers & Security*, 97:101968, 2020.

[92] P. Metzler, N. Suri, and G. Weissenbacher. Extracting safe thread schedules from incomplete model checking results. *International Journal on Software Tools for Technology Transfer*, 22(5):565–581, 2020.

[93] P. Mockapetris. Domain Names – Concepts and Facilities. RFC 1034, Oct. 1987.

[94] G. C. Moura, J. Heidemann, M. Müller, R. de O. Schmidt, and M. Davids. When the dike breaks: Dissecting dns defenses during ddos. In *Proceedings of the Internet Measurement Conference 2018*, pages 8–21, 2018.

[95] G. C. Moura, R. d. O. Schmidt, J. Heidemann, W. B. de Vries, M. Muller, L. Wei, and C. Hesselman. Anycast vs. ddos: Evaluating the november 2015 root dns event. In *Proceedings of the 2016 Internet Measurement Conference*, pages 255–270, 2016.

[96] G. C. M. Moura, S. Castro, W. Hardaker, M. Wullink, and C. Hesselman. Clouding up the internet: How centralized is dns traffic becoming? In *Proceedings of the ACM Internet Measurement Conference*, pages 42–49, 2020.

[97] N. Mukhtar, A. P. Fournaris, T. M. Khan, C. Dimopoulos, and Y. Kong. Improved hybrid approach for side-channel analysis using efficient convolutional neural network and dimensionality reduction. *IEEE Access*, 8:184298–184311, 2020.

[98] M. Müller, J. de Jong, M. van Heesch, B. Overeinder, and R. van Rijswijk-Deij. Retrofitting Post-Quantum Cryptography in Internet Protocols: a Case Study of DNSSEC. *ACM SIGCOMM Computer Communication Review*, 50(4):49–57, 2020.

[99] M. Müller, W. Toorop, T. Chung, J. Jansen, and R. van Rijswijk-Deij. The Reality of Algorithm Agility: Studying the DNSSEC Algorithm Life-Cycle. In *Proceedings of the ACM Internet Measurement Conference*, pages 295–308, 2020.

[100] P. Muth, M. Geihs, T. Arul, J. Buchmann, and S. Katzenbeisser. Elsa: efficient long-term secure storage of large datasets (full version). *EURASIP Journal on Information Security*, 2020:1–20, 2020.

[101] S. R. Niya, R. Beckmann, and B. Stiller. Dlit: A scalable distributed ledger for iot data. In *International Conference on Blockchain Computing and Applications (BCCA2020)*, Antalya, Turkey, November 2020. IEEE.

[102] S. R. Niya, D. Dordevic, and B. Stiller. Itrade: A blockchain-based, self-sovereign, and scalable marketplace for iot data streams. In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2021)*, Bordeaux, France, May 2021. IEEE.

[103] S. R. Niya, E. Schiller, I. Cepilov, and B. Stiller. BIIT: Standardization of Blockchain-based I2oT Systems in the I4 Era. In *Management in the Age of Softwarization and Artificial Intelligence*, pages 1–9, Budapest, Hungary, Apr. 2020. IEEE.

[104] R. Norvill, C. Cassanges, W. Shbair, J. Hilger, A. Cullen, and R. State. A security and privacy focused kyc data sharing platform. In *Proceedings of the 2nd ACM International Symposium on Blockchain and Secure Critical Infrastructure*, pages 151–160, 2020.

[105] D. Palma and T. Spatzier. Topology and orchestration specification for cloud applications (TOSCA). *Organization for the Advancement of Structured Information Standards (OASIS), Tech. Rep*, 2013.

[106] K. Papadamou, A. Papasavva, S. Zannettou, J. Blackburn, N. Kourtellis, I. Leontiadis, G. Stringhini, and M. Sirivianos. Disturbed YouTube for kids: Characterizing and detecting inappropriate videos targeting young children. In *Proceedings of the International AAAI Conference on Web and Social Media*, May 2020.

[107] K. Papadamou, S. Zannettou, J. Blackburn, E. D. Cristofaro, G. Stringhini, and M. Sirivianos. "it is just a flu": Assessing the effect of watch history on youtube's pseudoscientific video recommendations. *arXiv preprint: available at* https://arxiv.org/abs/2010.11638, 2020.

[108] K. Papadamou, S. Zannettou, J. Blackburn, E. D. Cristofaro, G. Stringhini, and M. Sirivianos. Understanding the incel community on youtube. *arXiv preprint: available at* https://arxiv.org/abs/2001.08293, 2020.

[109] E. Papadogiannaki, D. Deyannis, and S. Ioannidis. Head (er) hunter: Fast intrusion detection using packet metadata signatures. In *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1–6. IEEE, 2020.

[110] S. Pinto and N. Santos. Demystifying arm trustzone: A comprehensive survey. *ACM Computing Surveys (CSUR)*, 51(6):1–36, 2019.

[111] B. Podgorelec, M. Heričko, and M. Turkanović. State channel as a service based on a distributed and decentralized web. *IEEE Access*, 8:64678–64691, 2020.

[112] B. Podgorelec, M. Turkanović, and S. Karakatič. A machine learning-based method for automated blockchain transaction signing including personalized anomaly detection. *Sensors*, 20(1):147, 2020.

[113] B. Podgorelec, M. Turkanović, and M. Šestak. A brief review of database solutions used within blockchain platforms. In *International Congress on Blockchain and Applications*, volume 1238, pages 121–130. Springer, 2020.

[114] M. Poiitis, A. Vakali, and N. Kourtellis. On the aggression diffusion modeling and minimization in online social networks. *arXiv preprint arXiv:2005.10646*, 2020.

[115] R. Poschinger, N. Rodday, R. Labaca-Castro, and G. Dreo Rodosek. Openmtd: A framework for efficient network-level mtd evaluation. In *Proceedings of the 7th ACM Workshop on Moving Target Defense*, pages 31–41, 2020.

[116] S. Rafati Niya, B. Jeffrey, and B. Stiller. Kyot: Self-sovereign iot identification with a physically unclonable function. In *The 45th IEEE Conference on Local Computer Networks (LCN 2020)*, New York, NY, USA, November 2020. IEEE.

[117] S. Rafati Niya, E. Schiller, and B. Stiller. *Architectures for Blockchain-IoT Integration*. Wiley-IEEE Press, New York, NY, USA, December 2020.

[118] S. Rivera, V. K. Gurbani, S. Lagraa, A. K. Iannillo, and R. State. Leveraging ebpf to preserve user privacy for dns, dot, and doh queries. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*, pages 1–10, 2020.

[119] B. Rodrigues, E. Scheid, C. Killer, M. Franco, and B. Stiller. Blockchain signaling system (bloss): Cooperative signaling of distributed denial-of-service attacks. *Journal of Network and Systems Management*, 28(4):953–989, 2020.

[120] B. Rodrigues, E. J. Scheid, and B. Stiller. Blockchains in the age of softwarization - hands-on experiences with programming smart contracts and their security pitfalls, apr 2020.

[121] R. Roemer, E. Buchanan, H. Shacham, and S. Savage. Return-oriented programming: Systems, languages, and applications. *ACM Trans. Inf. Syst. Secur.*, 15(1), Mar. 2012.

[122] C. Rondanini, F. Daidone, B. Carminati, and E. Ferrari. Blockchain-based controlled information sharing in inter-organizational workflows. In *Proceeding of International Conference on Services Computing (SCC 2020)*, 2020.

[123] H. Saissi, S. Winter, O. Schwahn, K. Pattabiraman, and N. Suri. Tracesanitizer-eliminating the effects of non-determinism on error propagation analysis. In *2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 52–63. IEEE, 2020.

[124] N. Salamanos, M. J. Jensen, X. He, Y. Chen, C. Iordanou, and M. Sirivianos. Did state-sponsored trolls shape the us presidential election discourse? quantifying influence on twitter. *arXiv preprint: available at* https://arxiv.org/abs/2006.09938, 2020.

[125] B. Santos, B. Dzogovic, B. Feng, N. Jacot, T. Van Do, et al. Improving cellular iot security with identity federation and anomaly detection. In *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, pages 776–780. IEEE, 2020.

[126] E. Schiller, E. Esati, S. R. Niya, and B. Stiller. Blockchain on MSP430 with IEEE 802.15.4. In *Proceedings of the 45th IEEE Conference on Local Computer Networks (LCN'20)*, pages 1–4, Piscateway, New Jersey, U.S.A., Nov. 2020. IEEE.

[127] E. Schiller, S. Weber, and B. Stiller. Design and Evaluation of an SDR-based LoRa Cloud Radio Access Network. In *Proceedings of the 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'20)*, pages 1–7, Piscateway, New Jersey, U.S.A., Oct. 2020. IEEE.

[128] N. Schnepf, S. Merz, R. Badonnel, and A. Lahmadi. Automated Verification of Security Chains in Software-Defined Networks with Synaptic. In *Proceedings of the 3rd IEEE Conference on Network Softwarization (NetSoft'17)*, 2017.

[129] D. Serpanos, S. Yang, and M. Wolf. Neural network-based side channel attacks and countermeasures. In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pages 1–2, 2020.

[130] Y. Sheffer, R. Holz, and P. Saint-Andre. Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS). RFC 7525, May 2015.

[131] Y. Sheffer, R. Holz, and P. Saint-Andre. Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS). RFC 7525bis, Oct. 2020. https://tools.ietf.org/html/draft-ietf-uta-rfc7525bis-00.

[132] A. Shevtsov, M. Oikonomidou, D. Antonakaki, P. Pratikakis, and S. Ioannidis. Analysis of twitter and youtube during uselections 2020. *arXiv e-prints*, pages arXiv–2010, 2020.

[133] A. S. A. Shevtsov, M. Oikonomidou, D. Antonakaki, P. Pratikakis, A. Kanterakis, S. Ioannidis, and P. Fragopoulou. Discovery and classification of twitter bots. *arXiv preprint arXiv:2010.15393*, 2020.

[134] K. Solomos, P. Ilia, S. Ioannidis, and N. Kourtellis. Clash of the trackers: measuring the evolution of the online tracking ecosystem. In *IEEE/IFIP TMA Conference*, June 2020.

[135] R. Sommese, L. Bertholdo, G. Akiwate, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. Claffy, and A. Sperotto. Manycast2: Using anycast to measure anycast. In *Proceedings of the 2020 ACM Internet Measurement Conference*, pages 456–463. ACM, 2020.

[136] R. Sommese, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. C. Claffy, and A. Sperotto. The forgotten side of dns: Orphan and abandoned records. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, pages 538–543, 2020.

[137] R. Sommese, G. C. M. Moura, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. C. Claffy, and A. Sperotto. When Parents and Children Disagree: Diving into DNS Delegation Inconsistency. In *PAM 2020: Passive and Active Measurement*, pages 175–189. Springer, 2020.

[138] J. Steinberger, A. Sperotto, H. Baier, and A. Pras. Distributed ddos defense: A collaborative approach at internet scale. In *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6. IEEE, 2020.

[139] A. Taha, A. Zakaria, D. Kim, and N. Suri. Decentralized runtime monitoring approach relying on the ethereum blockchain infrastructure. In *2020 IEEE International Conference on Cloud Engineering (IC2E)*, pages 134–143. IEEE, 2020.

[140] C. Terizi, D. Chatzakou, E. Pitoura, P. Tsaparas, and N. Kourtellis. Angry birds flock together: Aggression propagation on social media. *arXiv preprint arXiv:2002.10131*, 2020.

[141] O. v. d. Toorn and A. Sperotto. Looking beyond the horizon: Thoughts on proactive detection of threats. *Digital Threats: Research and Practice*, 1(1), Mar. 2020.

[142] M. Tsantekidis and V. Prevelakis. Software System Exploration using Library Call Analysis. In *2nd Workshop on Model-driven Simulation and Training Environments for Cybersecurity (MSTEC)*, Sept. 2020.

[143] M. Turkanovic and B. Podgorelec. Signing blockchain transactions using qualified certificates. *IEEE Internet Computing*, 2020.

[144] O. v. der Toorn, R. van Rijswijk-Deij, T. Fiebig, M. Lindorfer, and A. Sperotto. Txting 101: Finding security issues in the long tail of dns txt records. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, pages 544–549, 2020.

[145] T. Wabeke, G. C. M. Moura, N. Franken, and C. Hesselman. Counterfighting counterfeit: Detecting and taking down fraudulent webshops at a cctld. In *Passive and Active Measurement*, pages 158–174, 2020.

[146] X. Zhang, J. Liu, J. Li, and L. Liu. Large-scale dynamic social network directed graph k-in&out-degree anonymity algorithm for protecting community structure. *IEEE Access*, pages 108371–108383, 2019.

[147] B. Z. H. Zhao, M. A. Kaafar, and N. Kourtellis. Not one but many Tradeoffs: Privacy Vs. Utility in Differentially Private Machine Learning. In *ACM Cloud Computing Security Workshop (CCSW)*, November 2020.

A Publications

The table on the next pages show all papers produced by CONCORDIA partners. The table shows the task to which the paper belongs, the partner's institute, the year of publication, the title and whether the paper is available as open access paper.

It should be noted that CONCORDIA has entered into the EU-ECAS system even more papers than what is shown in the list below. There are several reasons for that, including that white papers that were not peer-reviewed have been excluded from this Appendix, although they were found by the OpenAire system and therefore included into ECAS.

No	Task	Partner	Year	Title	Open Access
1	T1.1	JUB	2020	A Resource Efficient Implementation of the RESTCONF Protocol for OpenWrt Systems [60]	N
2	T1.1	FORTH	2020	An Enclave Assisted Snapshot-based Kernel Integrity Monitor [38]	N
3	T1.1	FORTH	2020	TrustAV: Practical and Privacy-Preserving Malware Analysis in the Cloud [39]	N
4	T1.1	BGU	2020	DANTE: A Framework for Mining and Monitoring Darknet Traffic [28]	N
5	T1.1	BGU	2020	When Explainability Meets Adversarial Learning: Detecting Adversarial Examples using SHAP Signatures [47]	N
6	T1.1	BGU	2020	IoT Device Identification Using Deep Learning [81]	N
7	T1.1	BGU	2020	A novel approach for detecting vulnerable IoT devices connected behind a home NAT [91]	N
8	T1.1	UMIL	2020	Requirements and Challenges for Secure and Trustworthy UAS Collaboration [11]	N
9	T1.1	UMIL	2020	Architectures for Blockchain-IoT Integration [117]	N
10	T1.1	UMIL	2020	KYoT: Self-Sovereign IoT Identification with A Physically Unclonable Function [116]	N
11	T1.1	UI	2021	Regulating data sharing across MQTT environments [32]	N
12	T1.1	UI	2021	Mammoth: Monitoring the ABAC Monitor of MQTT-Based Internet of Things Ecosystems [31]	Link
13	T1.1	UZH	2020	BIIT: Standardization of Blockchain-based I2oT Systems in the I4 Era [103]	Link
14	T1.1	UZH	2020	Blockchain on MSP430 with IEEE 802.15.4 [126]	Link
15	T1.1	UZH	2020	Design and Evaluation of an SDR-based LoRa Cloud Radio Access Network [127]	Link

Continued on next page . . .

No	Task	Partner	Year	Title	Open Access
16	T1.1	SIDN, UT	2020	The DNS in IoT: Opportunities, Risks, and Challenges [67]	N
17	T1.1	ISI	2020	Privacy Preservation in Industrial IoT via Fast Adaptive Correlation Matrix Completion [83]	Link
18	T1.1	FORTH	2020	An enclave assisted snapshot-based kernel integrity monitor [38]	Link
19	T1.1	UNI, PASSAU	2020	ELSA: efficient long-term secure storage of large datasets (full version) [100]	Link
20	T1.1	ISI, UP	2020	Anomaly Detection Trusted Hardware Sensors for Critical Infrastructure Legacy Devices [50]	Link
21	T1.1	UZH	2020	A Practical Analysis on Mirai Botnet Traffic [55]	Link
22	T1.1	ISI	2020	Decentralized, Secure and Cognitive Architecture for Automotive Cyber-Physical System of Systems [52]	Link
23	T1.1	ISI, CERTH	2020	Enabling Digital Forensics Readiness for Internet of Vehicles [7]	Link
24	T1.2	UT	2020	Global-Scale Anycast Network Management with Verfploeter [37]	Link
25	T1.2	UT	2020	A Longitudinal and Comprehensive Study of the {DANE} Ecosystem in Email [86]	Link
26	T1.2	UT	2020	TXTing 101: Finding Security Issues in the Long Tail of DNS TXT Records [144]	N
27	T1.2	UT	2020	The Forgotten Side of DNS: Orphan and Abandoned Records [136]	N
28	T1.2	UT	2020	When Parents and Children Disagree: Diving into DNS Delegation Inconsistency [137]	N
29	T1.2	UT	2020	Looking Beyond the Horizon: Thoughts on Proactive Detection of Threats [141]	Link
30	T1.2	UT	2020	Unresolved Issues: Prevalence, Persistence, and Perils of Lame Delegations [6]	Link

Continued on next page . . .

No	Task	Partner	Year	Title	Open Access
31	T1.2	UT	2020	MAnycast2: Using Anycast to Measure Anycast [135]	Link
32	T1.2	UT	2020	DDoS Mitigation: A Measurement-Based Approach [75]	N
33	T1.2	UT, SIDN	2020	The Reality of Algorithm Agility: Studying the DNSSEC Algorithm Life-Cycle [99]	N
34	T1.2	UT, SIDN	2020	Retrofitting Post-Quantum Cryptography in Internet Protocols: a Case Study of DNSSEC [98]	N
35	T1.2	UT, SIDN	2020	BGP Anycast Tuner: Intuitive Route Management for Anycast Services [17]	N
36	T1.2	FORTH	2020	Head (er) Hunter: Fast Intrusion Detection using Packet Metadata Signatures [109]	N
37	T1.2	UM	2020	State Channel as a Service Based on a Distributed and Decentralized Web [111]	Link
38	T1.2	SnT	2020	Leveraging eBPF to preserve user privacy for DNS, DoT, and DoH queries [118]	Link
39	T1.2	DFN-CERT	2020	ARIMA Supplemented Security Metrics for Quality Assurance and Situational Awareness [80]	Link
40	T1.2	UL	2020	Encrypted HTTP/2 Traffic Monitoring: Standing the Test of Time and Space [22]	Link
41	T1.2	MUNI	2020	Using TLS Fingerprints for OS Identification in Encrypted Traffic [84]	Link
42	T1.2	CODE, UT	2020	Towards Adversarial Resilience in Proactive Detection of Botnet Domain Names by using MTD [43]	Link
43	T1.2	UT	2020	Distributed ddos defense: A collaborative approach at internet scale [138]	Link
44	T1.2	UZH	2020	Blockchain Signaling System (BloSS): Cooperative Signaling of Distributed Denial-of-Service Attacks [119]	Link

Continued on next page ...

No	Task	Partner	Year	Title	Open Access
45	T1.2	UZH	2020	Design and implementation of cast-as-intended verifiability for a blockchain-based voting system [76]	Link
46	T1.2	UZH	2020	WeTrace–A Privacy-preserving Mobile COVID-19 Tracing Approach and Application [35]	Link
47	T1.2	UZH	2020	SC-FLARE: Cooperative DDoS Signaling based on Smart Contracts [35]	Link
48	T1.2	UZH	2020	Blockchains in the Age of Softwarization - Hands-on Experiences with Programming Smart Contracts and Their Security Pitfalls [120]	Link
49	T1.2	UZH	2020	Blockchains in the Age of Softwarization - Hands-on Experiences with Programming Smart Contracts and Their Security Pitfalls [120]	Link
50	T1.2	UZH	2020	ProtectDDoS: A Platform for Trustworthy Offering and Recommendation of Protections [53]	Link
51	T1.2	CODE	2020	OpenMTD: A Framework for Efficient Network-Level MTD Evaluation [115]	Link
52	T1.2	UZH	2020	SecBot: a Business-Driven Conversational Agent for Cybersecurity Planning and Management [54]	Link
53	T1.2	UZH	2020	Proverum: A Hybrid Public Verifiability and Decentralized Identity Management [79]	Link
54	T1.3	ULANC	2020	TraceSanitizer – Eliminating the Effects of Non-determinism on Error Propagation Analysis [123]	Link
55	T1.3	ULANC	2020	Decentralized Runtime Monitoring Approach Relying on the Ethereum Blockchain Infrastructure [139]	Link
56	T1.3	ULANC	2020	Extracting Safe Thread Schedules from Incomplete Model Checking Results [92]	Link

Continued on next page ...

No	Task	Partner	Year	Title	Open Access
57	T1.3	ULANC	2020	O Peer, Where Art Thou? Uncovering Remote Peering Interconnections at IXPs [59]	Link
58	T1.3	ULANC	2020	(How Much) Does a Private WAN Improve Cloud Performance? [15]	Link
59	T1.3	ULANC	2020	A First Look at the Misuse and Abuse of the IPv4 Transfer Market [58]	Link
60	T1.3	ULANC	2020	ICS Testbed Tetris: Practical Building Blocks Towards a Cyber Security Resource [61]	Link
61	T1.3	UL	2021	Compiler and optimization level recognition using graph neural networks [16]	Link
62	T1.3	SIDN, UT	2020	Clouding up the Internet: how centralized is DNS traffic becoming? [96]	N
63	T1.3	SIDN, UT	2020	Counterfighting Counterfeit: detecting and taking down fraudulent webshops at a ccTLD [145]	N
64	T1.3	UT, SIDN	2020	A Responsible Internet to Increase Trust in the Digital World [66]	Link
65	T1.3	UL	2021	Extracting Executable Payloads From Packed Malware: Import Table Reconstruction via Hardware-Assisted API Micro Execution [26]	N
66	T1.3	ULANC	2021	Decentralized Runtime Monitoring Approach Relying on the Ethereum Blockchain Infrastructure [139]	Link
67	T1.4	UM	2020	A Brief Review of Database Solutions Used within Blockchain Platforms [113]	N
68	T1.4	FORTH	2020	The Million Dollar Handshake: Secure and Attested Communications in the Cloud [25]	N
69	T1.4	FORTH	2020	Pythia: Scheduling of concurrent network packet processing applications on heterogeneous devices [57]	N
70	T1.4	UM	2020	A Brief Review of Database Solutions Used within Blockchain Platforms [113]	Link

Continued on next page ...

No	Task	Partner	Year	Title	Open Access
71	T1.4	UMIL	2020	Stay Thrifty, Stay Secure: A VPN-Based Assurance Framework for Hybrid Systems [10]	N
72	T1.4	UMIL	2020	A Blockchain-based Trustworthy Certification Process for Composite Services [12]	N
73	T1.4	UMIL	2020	A Methodology for Non-Functional Property Evaluation of Machine Learning Models [13]	N
74	T1.4	UMIL	2020	Be Your Neighbor's Miner: Building Trust in Ledger Content via Reciprocally Useful Work [87]	N
75	T1.4	TUBS	2020	Software System Exploration using Library Call Analysis [142]	Link
76	T1.4	TUBS	2020	SAVTA: A Hybrid Vehicular Threat Model: Overview and Case Study [62]	Link
77	T1.4	TUBS	2021	SPPS: Secure Policy-based Publish/Subscribe System for V2C Communication (Accepted) [63]	
78	T1.5	UI	2021	Evaluating the effects of access control policies within NoSQL systems [30]	Link
79	T1.5	UI	2020	A Blockchain-based Trustworthy Certification Process for Composite Services [12]	N
80	T1.4	UI	2020	Certified Machine-Learning Models [34]	Link
81	T1.4	OsloMET / Telenor	2020	Secure healthcare: 5G-enabled network slicing for elderly care [45]	Link
82	T1.4	OsloMET / Telenor	2020	Improving Cellular IoT Security with Identity Federation and Anomaly Detection [125]	Link
83	T1.4	UI	2020	Blockchain-based controlled information sharing in inter-organizational workflows [122]	N
84	T1.5	FORTH	2020	The Seven Deadly Sins of the HTML5 WebAPI: A Large-scale Study on the Risks of Mobile Sensor-based Attacks [40]	N

Continued on next page . . .

No	Task	Partner	Year	Title	Open Access
85	T1.5	FORTH	2020	Analysis of Twitter and YouTube during USelections 2020 [132]	Link
86	T1.5	FORTH	2020	Discovery and classification of Twitter bots [133]	Link
87	T1.5	UM	2020	A Machine Learning-Based Method for Automated Blockchain Transaction Signing Including Personalized Anomaly Detection [112]	N Link
88	T1.5	UM	2020	Signing Blockchain Transactions using Qualified Certificates [143]	Link
89	T1.5	SnT	2020	A Security and Privacy Focused KYC Data Sharing Platform [104]	Link
90	T1.5	TID, CUT	2021	Catching them red-handed: Real-time Aggression Detection on Social Media[65]	Link
91	T1.5	TID, CUT	2020	A Streaming Machine Learning Framework for Online Aggression Detection on Twitter[64]	Link
92	T1.5	TID, CUT	2020	Disturbed YouTube for kids: Characterizing and detecting inappropriate videos targeting young children[106]	Link
93	T1.5	TID	2020	Not one but many Tradeoffs: Privacy Vs. Utility in Differentially Private Machine Learning[147]	Link
94	T1.5	TID	2020	FLaaS: Federated Learning as a Service[82]	Link
95	T1.5	TID, FORTH	2020	Clash of the trackers: measuring the evolution of the online tracking ecosystem[134]	Link
96	T1.5	TID	2020	Stop tracking me Bro! Differential Tracking of User Demographics on Hyper-Partisan Websites[5]	Link
97	T1.5	TID	2020	ITrade: A Blockchain-based, Self-Sovereign, and Scalable Marketplace for IoT Data Streams [102]	N
98	T1.5	TID	2020	DLIT: A Scalable Distributed Ledger for IoT Data [101]	N
99	T1.5	CUT	2020	Did State-sponsored Trolls Shape the US Presidential Election Discourse? Quantifying Influence on Twitter [124]	Link

Continued on next page . . .

No	Task	Partner	Year	Title	Open Access
100	T1.5	CUT	2020	Understanding the Incel Community on YouTube [108]	Link
101	T1.5	CUT	2020	"It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations [107]	Link
102	T1.5	UI	2021	Privacy-Preserving Sequentially Publishing of Knowledge Graphs [70]	N
103	T1.5	UI	2020	Cluster-based Anonymization of Knowledge Graphs [69]	Link
104	T1.5	UI	2019	Cluster-based Anonymization of Directed Graphs [68]	Link
105	T1.5	TID	2020	Angry Birds Flock Together: Aggression Propagation on Social Media [8]	Link
106	T1.5	UI	2020	A deep learning model for Twitter spam detection [140]	Link
107	T1.5	TID	2020	On the Aggression Diffusion Modeling and Minimization in Online Social Networks [114]	Link

B Organization of Conferences

The table on the next pages shows all conferences that are or have been organized by CONCORDIA partners. Organization implies one of the following roles: General (co)chair, TPC (co)chair, Program (co)chair, Local chair or Organizing Committee (co)chair. The table is organized in alphabetic order of the last name.

Scientific Event	Abbr.	Where	When	URL
10th ACM Conference on Data and Application Security and Privacy	**ACM CODASPY 2020**	**Virtual Conference**	**3-4 Aug 2020**	http://www.codaspy.org/2020/
Name: Barbara Carminati	Partner:	UI	Role:	Program Co-Chair
11th ACM Conference on Data and Application Security and Privacy	**ACM CODASPY 2021**	**Virtual Conference**	**22-24 Mar 2021**	http://www.codaspy.org/2021/
Name: Barbara Carminati	Partner:	UI	Role:	Program Co-Chair
Name: Elena Ferrari	Partner:	UI	Role:	Workshop Chair
12th IEEE International Conference on Cloud Computing	**IEEE CLOUD 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/cloud/2020/
Name: Claudio Ardagna	Partner:	UMIL	Role:	Program Co-Chair
12th IEEE International Conference on Web Services	**IEEE ICWS 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/icws/2020/
Name: Elena Ferrari	Partner:	UI	Role:	Program Co-Chair
16th International Conference on emerging Networking EXperiments and Technologies	**CoNEXT 2020**	**Virtual Conference**	**1-4 Dec 2020**	https://conferences2.sigcomm.org/co-next/2020/

Continued on next page...

Scientific Event		Abbr.	Where	When	URL
Name:	Ralph Holz	Partner:	UT	Role:	Reproducibility Co-Chair & Artifact Evaluation Committee
Name:	Nicolas Kourtellis	Partner:	TID	Role:	Local Co-Chair
35th International Conference on ICT Systems Security and Privacy Protection		**IFIP SEC 2020**	**Virtual Conference**	**21-23 Sep 2020**	https://sec2020.um.si/
Name:	Tatjana Welzer	Partner:	UM	Role:	General chair
Name:	Lili Nemec Zlatolas	Partner:	UM	Role:	Organizing chair
36th IEEE International Conference on Data Engineering		**IEEE ICDE 2020**	**Virtual Conference**	**20-24 Apr 2020**	https://www.utdallas.edu/icde/
Name:	Barbara Carminati	Partner:	UI	Role:	Research PC Vice Chair
Name:	Elena Ferrari	Partner:	UI	Role:	General Co-Chair
6th IEEE/IFIP Workshop on Security for Emerging Distributed Network Technologies		**DISSECT 2020**	**Virtual Workshop**	**24 Apr 2020**	http://www.inf.ufrgs.br/dissect/2020/
Name:	Thibault Cholez	Partner:	UL	Role:	Organizing Committee
ACM/IRTF Applied Networking Research Workshop		**ANRW 2020**	**Virtual Workshop**	**30-31 Jul 2020**	https://irtf.org/anrw/2020/
Name:	Roland van Rijswijk	Partner:	UT	Role:	Program Co-Chair

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
IEEE 19th International Conference on Industrial Informatics - Track 3. Safety and Security in Industrial Applications	**INDIN 2021**	**Palma de Mallorca, Spain**	**21-23 Jul 2021**	https://2021.ieee-indin.org/technical-technical_tracks/
Name: Dimitrios Serpanos	Partner:	ISI	Role:	Track Co-Chair
IEEE Services Doctoral Symposium 2020	**IEEE DS SERVICES 2020**	**Virtual Conference**	**20-24 Oct 2020**	https://conferences.computer.org/services/2020/symposia/ds.html
Name: Barbara Carminati	Partner:	UI	Role:	Committee Co-Chair
Passive and Active Measurement conference 2020	**PAM 2020**	**Virtual Conference**	**30-31 Mar 2020**	https://pam2020.cs.uoregon.edu/
Name: Anna Sperotto	Partner:	UT	Role:	Program Co-Chair
The 13th International Symposium on Foundations and Practice of Security	**FPS 2020**	**Montreal, Canada**	**1-3 Dec 2020**	https://www.fps-symposium.com/home
Name: Jean-Yves Marion	Partner:	UL	Role:	General Co-Chair
The Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications	**TPS 2020**	**Virtual Conference**	**1-3 Dec 2020**	http://www.sis.pitt.edu/lersais/conference/tps/2020/

Continued on next page...

Scientific Event		Abbr.	Where	When	URL
Name:	Barbara Carminati	Partner:	UI	Role:	Program Co-Chair

C Technical Program Committee Membership

The table on the next pages shows all conferences for which CONCORDIA partners are or have been member of the Technical Program Committee (TPC). The table is organized in alphabetic order of the first name.

Scientific Event	Abbr.	Where	When	URL
10th ACM Conference on Data and Application Security and Privacy	**ACM CODASPY 2020**	**Virtual Conference**	**3-4 Aug 2020**	http://www.codaspy.org/2020/
Name: Pietro Colombo			Partner: UI	
Name: Elena Ferrari			Partner: UI	
12th ACM Web Science Conference 2020	**WebSci 20**	**Virtual Conference**	**6-10 Jul 2020**	https://websci20.webscience.org
Name: Nicolas Kourtellis			Partner: TID	
13th IEEE/ACM International Conference on Utility and Cloud Computing	**UCC 2020**	**Leicester, UK**	**7-10 Dec 2020**	https://www.cs.le.ac.uk/events/UCC2020/
Name: Claudio Ardagna			Partner: UMIL	
14th International Baltic Conference on Databases and Information Systems	**Baltic DB&IS 2020**	**Virtual Conference**	**16-19 Jun 2020**	https://dbis.ttu.ee/index.php?page=65
Name: Tatjana Welzer			Partner: UM	
14th International Conference on Web and Social Media	**ICWSM 2020**	**Virtual Conference**	**8-11 Jun 2020**	https://www.icwsm.org/2020/index.html
Name: Nicolas Kourtellis			Partner: TID	
Name: Michael Sirivianos			Partner: CUT	

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
15th Dependable, Adaptive, and Secure Distributed Systems	**DADS 2020**	**Brno, Czech Republic**	**30 Mar-3 Apr 2020**	https://www.dedisys.org/sac20/
Name:	Claudio Ardagna		Partner:	UMIL
16th International Conference on Information Systems Security	**ICISS 2020**	**Virtual Conference**	**16-20 Dec 2020**	https://isrdc.iitb.ac.in/iciss2020/
Name:	Claudio Ardagna		Partner:	UMIL
16th International Conference on Network and Service Management	**CNSM 2020**	**Virtual Conference**	**2-6 Nov 2020**	http://www.cnsm-conf.org/2020/
Name:	Mattijs Jonker		Partner:	UT
Name:	Jürgen Schönwälder		Partner:	JUB
Name:	Anna Sperotto		Partner:	UT
1st Workshop on Conceptual Modeling Meets Artificial Intelligence and Data-Driven Decision Making	**CMAI 2020**	**Virtual Conference**	**3-6 Nov 2020**	https://workshop-cmai.github.io/2020/
Name:	Tatjana Welzer		Partner:	UM
2020 IEEE International Conference on Big Data	**IEEE BigData 2020**	**Virtual Conference**	**10-13 Dec 2020**	https://bigdataieee.org/BigData2020/
Name:	Pietro Colombo		Partner:	UI

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
2020 IEEE International Conference on Services Computing	**IEEE SCC 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/scc/2020/
Name:	Pietro Colombo		Partner:	UI
2020 IEEE International Conference on Web Services	**IEEE ICWS 2020**	**Virtual Conference**	**18-24 October 2020**	https://conferences.computer.org/icws/2020/
Name:	Marco Anisetti		Partner:	UMIL
Name:	Pietro Colombo		Partner:	UI
Name:	Barbara Carminati		Partner:	UI
2020 IEEE International Symposium on Smart Electronic Systems (iSES)	**iSES 2020**	**Virtual Conference**	**14-16 Dec 2020**	http://www.ieee-ises.org/
Name:	Apostolos Fournaris		Partner:	ISI
2020 IEEE/IFIP Network Operations and Management Symposium	**NOMS 2020**	**Virtual Conference**	**20-24 Apr 2020**	https://noms2020.ieee-noms.org/
Name:	Thibault Cholez		Partner:	UL
Name:	Cristian Hesselman		Partner:	SIDN
Name:	Jürgen Schönwälder		Partner:	JUB
Name:	Anna Sperotto		Partner:	UT
Name:	Pavel Čeleda		Partner:	MUNI
Name:	Martin Drašar		Partner:	MUNI

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
2021 IFIP/IEEE International Symposium on Integrated Network Management	**IM 2021**	**Bordeaux, France**	**17-21 May 2021**	https://im2021.ieee-im.org/
Name: Thibault Cholez			Partner: UL	
Name: Pavel Čeleda			Partner: MUNI	
Name: Martin Drašar			Partner: MUNI	
24th International Conference on Extending Database Technology	**EDBT 2021**	**Nicosia, Cyprus**	**23-26 Mar 2021**	https://edbticdt2021.cs.ucy.ac.cy/
Name: Elena Ferrari			Partner: UI	
25th IEEE Pacific Rim International Symposium on Dependable Computing	**PRDC 2020**	**Perth, Australia**	**1-4 Dec 2021**	http://prdc.dependability.org/PRDC2020/
Name: Neeraj Suri			Partner: ULANC	
25th IEEE Symposium on Computers and Communications	**ISCC 2020**	**Rennes, France**	**7-10 Jul 2020**	http://conferences.imt-atlantique.fr/iscc2020/
Name: Dimitrios Serpanos			Partner: ISI	
27th International Conference on Telecommunications	**ICT 2020**	**Virtual Conference**	**5-7 Oct 2020**	https://ict-20.org/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Dimitrios Serpanos		Partner:	ISI
28th Italian symposium on advanced database systems	**SEBD 2020**	**Virtual Conference**	**21-24 Jun 2020**	https://sebd2020.unica.it/
Name:	Elena Ferrari		Partner:	UI
29th ACM International Conference on Information and Knowledge Management	**CIKM 2020**	**Virtual Conference**	**19-23 Oct 2020**	https://cikm2020.org
Name:	Nicolas Kourtellis		Partner:	TID
2nd conference on Blockchain Research & Applications for Innovative Networks and Services	**BRAINS 2020**	**Virtual Conference**	**28-30 Sep 2020**	https://brains.dnac.org/
Name:	Thibault Cholez		Partner:	UL
2nd Workshop on Attackers and Cyber-Crime Operations	**WACCO 2020**	**Virtual Workshop**	**7 Sep 2020**	https://www.wacco-workshop.eu/
Name:	Ralph Holz		Partner:	UT
35th International Conference on ICT Systems Security and Privacy Protection	**IFIP SEC 2020**	**Virtual Conference**	**21-23 Sep 2020**	https://sec2020.um.si/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name: Muhamed Turkanović Name: Lili Nemec Zlatolas			Partner: UM Partner: UM	
39th International Symposium on Reliable Distributed Systems	**SRDS 2020**	**Shanghai, China**	**21-24 Sep 2020**	https://srds-conference.org/index-real.html#detail-notice
Name: Neeraj Suri			Partner: ULANC	
45th IEEE Conference on Local Computer Networks	**LCN 2020**	**Virtual Conference**	**16-19 Nov 2020**	https://www.ieeelcn.org/
Name: Dimitrios Serpanos			Partner: ISI	
50th IEEE/IFIP Int. Conference on Dependable Systems and Networks	**DSN 2020**	**Valencia, Spain**	**29 Jun - 2 Jul 2020**	https://dsn2020.webs.upv.es/
Name: Neeraj Suri			Partner: ULANC	
5th IEEE/IFIP International Workshop on Analytics for Network and Service Management	**AnNet 2020**	**Virtual Workshop**	**20 Apr 2020**	https://annet2020.loria.fr/
Name: Anna Sperotto			Partner: UT	

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
5th International Workshop on Traffic Measurements for Cybersecurity	**WTMC 2020**	**Virtual Workshop**	**7 Sep 2020**	https://wtmc.info/index2020.html
Name:	Ralph Holz		Partner:	UT
Name:	Roland van Rijswijk		Partner:	UT
Name:	Anna Sperotto		Partner:	UT
6th International Conference on Information Systems Security and Privacy	**ICISSP 2020**	**Valletta, Malta**	**25-27 Feb 2020**	http://www.icissp.org/?y=2020
Name:	Lara Mauri		Partner:	UMIL
Name:	Apostolos Fournaris		Partner:	ISI
6th International Workshop on Traffic Measurements for Cybersecurity	**WTMC 2021**	**Virtual Workshop**	**27 May 2021**	https://wtmc.info/index.html
Name:	Christian Keil		Partner:	DFN-CERT
ACM International Conference on Knowledge Discovery and Data Mining	**KDD 2020**	**Virtual Conference**	**23-27 Aug 2020**	https://www.kdd.org/kdd2020/
Name:	Nicolas Kourtellis		Partner:	TID
ACM International Conference on Management of Data	**ACM SIGMOD 2021**	**Xian, China**	**20-25 Jun 2021**	https://2021.sigmod.org/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Elena Ferrari		Partner:	UI
ACM SIGMETRICS 2020	**SIGMETRICS 2020**	**Virtual Conference**	**8-12 Jun 2020**	http://www.sigmetrics.org/sigmetrics2020/
Name:	Michael Sirivianos		Partner:	CUT
ACM Symposium on Access Control Model Technologies	**ACM SACMAT 2020**	**Virtual Conference**	**10-12 Jun 2020**	http://www.sacmat.org/2020/index.php
Name:	Barbara Carminati		Partner:	UI
Name:	Elena Ferrari		Partner:	UI
ACM/IRTF Applied Networking Research Workshop	**ANRW 2020**	**Virtual Workshop**	**30-31 Jul 2020**	https://irtf.org/anrw/2020/
Name:	Ralph Holz		Partner:	UT
Name:	Roland van Rijswijk		Partner:	UT
Design Automation Conference 2020	**DAC 2020**	**San Francisco**	**11-15 Jul 2020**	https://www.dac.com/
Name:	Dimitrios Serpanos		Partner:	ISI
Digital Forensics Research Workshop EU 2021	**DFRWS EU 2021**	**Virtual Conference**	**29 Mar - 1 Apr 2021**	https://dfrws.org/conferences/dfrws-eu-2021/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Christian Keil		Partner:	DFN-CERT
Eighth International Symposium on Security in Computing and Communications	**SSCC 2020**	**Chennai, India**	**14-17 Oct 2020**	http://www.acn-conference.org/2020/sscc2020/
Name:	Claudio Ardagna		Partner:	UMIL
Name:	Apostolos Fournaris		Partner:	ISI
Euromicro Conference on Digital System Design	**DSD 2020**	**Virtual Conference**	**26-28 Aug 2020**	https://dsd-seaa2020.um.si/dsd/
Name:	Apostolos Fournaris		Partner:	ISI
European Conference on Advances in Databases and Information Systems	**ADBIS 2020**	**Virtual Conference**	**25-27 Aug 2020**	http://eric.univ-lyon2.fr/adbis-tpdl-eda-2020/adbis/committee-adbis/
Name:	Tatjana Welzer		Partner:	UM
IEEE International Conference on Blockchain and Cryptocurrency 2020	**ICBC 2020**	**Virtual Conference**	**2-6 May 2020**	https://icbc2020.ieee-icbc.org/
Name:	Thibault Cholez		Partner:	UL
Name:	Wazen Shbair		Partner:	SnT-UL

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
IEEE International Conference on Blockchain and Cryptocurrency 2021	**ICBC 2021**	**Virtual Conference**	**3-6 May 2021**	https://icbc2021.ieee-icbc.org/
Name:	Wazen Shbair		Partner:	SnT-UL
IEEE International Conference on Cloud Computing	**IEEE CLOUD 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/cloud/2020/
Name:	Elena Ferrari		Partner:	UI
IEEE International Conference on Cloud Networking	**CloudNet 2020**	**Virtual Conference**	**9-11 Nov 2020**	https://cloudnet2020.ieee-cloudnet.org/
Name:	Thibault Cholez		Partner:	UL
IEEE International Conference on Edge Computing	**IEEE EDGE 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/edge/2020/
Name:	Marco Anisetti		Partner:	UMIL
IEEE International Conference on Omni-layer Intelligent Systems	**COINS 2020**	**Virtual Conference**	**31 Aug - 2 Sep 2020**	https://coinsconf.com/
Name:	Wazen Shbair		Partner:	SnT-UL
IEEE International Conference on Smart Data Services	**SMDS 2020**	**Virtual Conference**	**18-24 Oct 2020**	https://conferences.computer.org/smds/2020/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Claudio Ardagna		Partner:	UMIL
IFIP Networking 2020 Conference	**IFIP Networking 2020**	**Virtual Conference**	**22-25 Jun 2020**	https://networking.ifip.org/2020/
Name:	Jürgen Schönwälder		Partner:	JUB
International Conference on Information Technology and Communications Security 2020	**SecITC 2020**	**Virtual Conference**	**19-20 Nov 2020**	https://www.secitc.eu/
Name:	Claudio Ardagna		Partner:	UMIL
International Conference on INnovations in Intelligent SysTems and Applications	**INISTA 2020**	**Virtual Conference**	**24-26 Aug 2020**	http://inista.org/program-committee.php
Name:	Tatjana Welzer		Partner:	UM
International Conference on Security and Privacy in Digital Economy	**SPDE 2020**	**Quzhou, China**	**30 Oct - 1 Nov 2020**	http://spde2020.csp.escience.cn/
Name:	Claudio Ardagna		Partner:	UMIL
Italian conference on Cybersecurity	**ITASEC 2020**	**Ancona, Italy**	**4-7 Feb 2020**	https://itasec.it/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Elena Ferrari		Partner:	UI
Network Traffic Measurement and Analysis Conference	**TMA 2020**	**Virtual Conference**	**10-11 Jun 2020**	https://tma.ifip.org/2020
Name:	Ralph Holz		Partner:	UT
Name:	Roland van Rijswijk		Partner:	UT
Passive and Active Measurement conference 2020	**PAM 2020**	**Virtual Conference**	**30-31 Mar 2020**	https://pam2020.cs.uoregon.edu/
Name:	Cristian Hesselman		Partner:	SIDN
Name:	Ralph Holz		Partner:	UT
Name:	Roland van Rijswijk		Partner:	UT
The 12th IEEE International Conference on Cloud Computing Technology and Science	**CloudCom 2020**	**Bangkok, Thailand**	**14-17 Dec 2020**	https://2020.cloudcom.org/
Name:	Claudio Ardagna		Partner:	UMIL
The 12th International Congress on Ultra Modern Telecommunications and Control Systems	**ICUMT 2020**	**Virtual Conference**	**5-7 Oct 2020**	https://icumt.info/2020/
Name:	Dimitrios Serpanos		Partner:	ISI

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
The 15th International Conference on Availability, Reliability and Security	**ARES 2020**	**Virtual Conference**	**25-28 Aug 2020**	https://www.ares-conference.eu/conference-2020/cfp2020/
Name:	Marco Cremonini		Partner:	UMIL
The 16th International Conference on emerging Networking EXperiments and Technologies	**CoNEXT 2020**	**Virtual Conference**	**1-4 Dec 2020**	https://conferences2.sigcomm.org/co-next/2020/
Name:	Michael Sirivianos		Partner:	CUT
The 16th International Conference on emerging Networking EXperiments and Technologies – Artifact Evaluation	**CoNEXT 2020 (AEC)**	**Virtual Conference**	**1-4 Dec 2020**	https://conferences2.sigcomm.org/co-next/2020/
Name:	Mattijs Jonker		Partner:	UT
Name:	Roland van Rijswijk		Partner:	UT
Name:	Anna Sperotto		Partner:	UT
The 19th IEEE International Conference on Trust, Security and Privacy in Computing and Communications	**TrustCom 2020**	**Guangzhou, China**	**29 Dec 2020 - 1 Jan 2021**	http://ieee-trustcom.org/TrustCom2020/
Name:	Claudio Ardagna		Partner:	UMIL

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
The 2020 ACM Special Interest Group on Data Communication – Artifact Evaluation	**SIGCOMM 2020 (AEC)**	**Virtual Conference**	**11-13 Aug 2020**	https://conferences.sigcomm.org/sigcomm/2020/
Name:	Roland van Rijswijk		Partner:	UT
The 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining	**ASONAM 2020**	**Virtual Conference**	**7-10 Dec 2020**	http://asonam.cpsc.ucalgary.ca/2020/
Name:	Elena Ferrari		Partner:	UI
The 28th IEEE International Conference on Network Protocols	**ICNP 2020**	**Virtual Conference**	**13-16 Oct 2020**	https://icnp20.cs.ucr.edu/
Name:	Ralph Holz		Partner:	UT
The 6th IEEE International Conference on Big Data Computing Service and Machine Learning Applications	**IEEE BigDataService 2020**	**Oxford, United Kingdom**	**13-16 Apr 2020**	http://www.big-dataservice.net/
Name:	Pietro Colombo		Partner:	UI
The 6th IEEE International Conference on Collaboration and Internet Computing	**IEEE CIC 2020**	**Virtual Conference**	**1-3 Dec 2020**	http://www.sis.pitt.edu/lersais/conference/cic/2020/

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Elena Ferrari		Partner:	UI
The 8th International Conference on Future Internet of Things and Cloud	**FiCloud 2020**	**Virtual Conference**	**24-26 Aug 2020**	http://www.ficloud.org/2020/
Name:	Marco Anisetti		Partner:	UMIL
Name:	Dimitrios Serpanos		Partner:	ISI
The Fifth Workshop on Computational Methods in Online Misbehavior	**CyberSafety 2020**	**Virtual Workshop**	**21 Apr 2020**	https://cybersafety-workshop.github.io/2020/
Name:	Nicolas Kourtellis		Partner:	TID
The Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications	**IEEE TPS 2020**	**Virtual Conference**	**1-3 Dec 2020**	http://www.sis.pitt.edu/lersais/conference/tps/2020/
Name:	Elena Ferrari		Partner:	UI
The Second International Workshop on Blockchain Applications and Theory	**BAT 2020**	**Virtual Event**	**30 Jun - 3 Jul 2020**	http://emergingtechnet.org/BAT2020/
Name:	Elena Ferrari		Partner:	UI
The Web Conference 2021	**WWW 2021**	**Ljubljana, Slovenia**	**19-23 Apr 2021**	https://www2021.thewebconf.org

Continued on next page...

Scientific Event	Abbr.	Where	When	URL
Name:	Neeraj Suri		Partner:	ULANC
Name:	Barbara Carminati		Partner:	UI
Name:	Elena Ferrari		Partner:	UI
Usenix Eurosys	**EuroSys 2021**	**Edinburgh, UK**	**26-29 Apr 2021**	https://2021.eurosys.org/
Name:	Neeraj Suri		Partner:	ULANC
Workshop on Cyber Range Technologies and Applications	**CACOE'20**	**Virtual Workshop**	**7 Sep 2020**	https://cacoe.ait.ac.at/
Name:	Pavel Čeleda		Partner:	MUNI

D Editors of Journals

The table on the next pages shows all journals for which CONCORDIA members act as editors. Editor roles can be: Editor in Chief, Series Editor, Associate Editor, Area Editor, Guest Editor, Editorial Board Member or Editorial Advisory Board Member.

Description	Publisher	URL
ACM DTRAP – Special Issue on Vulnerabilities	**ACM**	https://dtrap-blog.acm.org/2020/08/28/special-issue-on-vulnerabilities/
Name: Jeroen van der Ham	Partner: UT	Role: Editorial board member
ACM Transactions on Data Science	**ACM**	https://tds.acm.org/editorial.cfm
Name: Elena Ferrari	Partner: UI	Role: Associate Editor
ACM Transactions on Privacy and Security	**ACM**	https://dl.acm.org/journal/tops
Name: Elena Ferrari	Partner: UI	Role: Associate Editor
Communications Magazine	**IEEE**	https://www.comsoc.org/publications/magazines/ieee-communications-magazine
Name: Jürgen Schönwälder	Partner: JUB	Role: Series Editor
Computer Communications Review	**ACM**	https://ccronline.sigcomm.org/
Name: Anna Sperotto	Partner: UT	Role: Editorial board member
Cybersecurity and Privacy of Frontiers in Big Data	**Frontiers Media SA**	https://www.frontiersin.org/journals/big-data/sections/cybersecurity-and-privacy
Name: Elena Ferrari	Partner: UI	Role: Specialty Chief Editor

Continued on next page...

Description	Publisher	URL
Data Science and Engineering	**Springer**	https://www.springer.com/journal/41019
Name: Elena Ferrari	Partner: UI	Role: Associate Editor
Digital Threats: Research and Practice	**ACM**	https://dl.acm.org/journal/dtrap
Name: Jeroen van der Ham	Partner: UT	Role: Editorial board member
Digital Trust: Trust Management in the Cyberspace, IEEE Internet computing	**IEEE**	https://www.computer.org/digital-library/magazines/ic/call-for-papers-special-issue-on-digital-trust-trust-management-in-the-cyberspace
Name: Elena Ferrari	Partner: UI	Role: Special Issue Editor
IEEE Internet Computing	**IEEE**	https://www.computer.org/csdl/magazine/ic/about/15624?title=Editorial%20Board&periodical=IEEE%20Internet%20Computing
Name: Elena Ferrari	Partner: UI	Role: Associate Editor in Chief
IEEE Transactions on Big Data	**IEEE**	https://ieeexplore.ieee.org/xpl/aboutJournal.jsp?punumber=6687317
Name: Neeraj Suri	Partner: ULANC	Role: Associate Editor

Continued on next page...

Description	Publisher	URL
IEEE Transactions on Cloud Computing	**IEEE**	https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6245519
Name: Neeraj Suri	Partner: ULANC	Role: Associate Editor
IEEE Transactions on Emerging Topics in Computing	**IEEE**	https://www.computer.org/csdl/journal/ec
Name: Dimitrios Serpanos	Partner: ISI	Role: Associate Editor
IEEE Transactions on Information Forensics and Security	**IEEE**	https://signalprocessingsociety.org/publications-resources/ieee-transactions-information-forensics-and-security/about-transactions
Name: Emil Lupu	Partner: Imperial	Role: Associate Editor
IEEE Transactions on Network and Service Management	**IEEE**	https://www.comsoc.org/publications/journals/ieee-tnsm
Name: Jürgen Schönwälder	Partner: JUB	Role: Editorial board member
IEEE Transactions on Service Computing	**IEEE**	https://www.computer.org/csdl/journal/sc/misc/14407?title=About&periodical=IEEE%20Transactions%20on%20Services%20Computing
Name: Elena Ferrari	Partner: UI	Role: Associate Editor

Continued on next page...

Description	Publisher	URL
Information Security Journal: A Global Perspective	**Taylor and Francis**	https://www.tandfonline.com/toc/uiss20/current
Name: Apostolos Fournaris	Partner: ISI	Role: Editorial board member/Associate Editor
International Journal of Cooperative Information Systems	**World Scientific**	https://www.worldscientific.com/page/ijcis/editorial-board
Name: Elena Ferrari	Partner: UI	Role: Associate Editor
International Journal of Network Management	**Wiley**	https://onlinelibrary.wiley.com/journal/10991190
Name: Jürgen Schönwälder	Partner: JUB	Role: Editorial board member
Name: Anna Sperotto	Partner: UT	Role: Editorial board member
Name: Emil Lupu	Partner: Imperial	Role: Editorial board member
Journal of Network and Systems Management	**Springer**	https://www.springer.com/journal/10922
Name: Cristian Hesselman	Partner: SIDN	Role: Co-guest Editor
Name: Jürgen Schönwälder	Partner: JUB	Role: Editorial board member
Name: Anna Sperotto	Partner: UT	Role: Editorial board member
Name: Emil Lupu	Partner: Imperial	Role: Editorial board member
Secure, Efficient Cyber-Physical Systems and Wireless Sensors	**MDPI**	https://www.mdpi.com/journal/jsan/special_issues/Cyber_physical_Sensors

Continued on next page...

Description	Publisher	URL
Name: Apostolos Fournaris	Partner: ISI	Role: Special Issue Editor
Sensors	**MDPI**	https://www.mdpi.com/journal/sensors/special_issues/Blockchain_Identity_Management_ICT_IoT
Name: Muhamed Turkanović	Partner: UM	Role: Special Issue Editor
Special issue on Computing for Autonomy: Latency, Power, Resilience	**IEEE Computer**	https://www.computer.org/digital-library/magazines/co/call-for-papers-special-issue-on-computing-for-autonomy-latency-power-resilience
Name: Dimitrios Serpanos	Partner: ISI	Role: Guest editor
Sustainability	**MDPI**	https://www.mdpi.com/journal/sustainability/special_issues/dte_sus
Name: Muhamed Turkanović	Partner: UM	Role: Special Issue Editor
Synthesis Lectures on Security, Privacy and Trust	**Morgan and Claypool**	https://www.morganclaypool.com/toc/spt/1/1
Name: Elena Ferrari	Partner: UI	Role: Editor