RESEARCH CENTRE

Inria Centre at Rennes University

IN PARTNERSHIP WITH:

CentraleSupélec, CNRS, Université de Rennes

2024

ACTIVITY REPORT

Project-Team

PIRAT

Protection of Information and Resistance to ATtacks

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

DOMAIN

Algorithmics, Programming, Software and Architecture

THEME

Security and Confidentiality

Inria

Contents

Project-Team PIRAT

Creation of the Project-Team: 2024 March 01

Keywords

Computer sciences and digital sciences

A1.2.2. – Supervision
A1.2.8. – Network security
A1.3.3. – Blockchain
A1.3.4. – Peer to peer
A2.2.1. – Static analysis
A4.1. – Threat analysis
A4.4. – Security of equipment and software
A4.9. – Security supervision
A4.9.1. – Intrusion detection
A4.9.2. – Alert correlation
A4.9.3. – Reaction to attacks
A9.1. – Knowledge
A9.2. – Machine learning
A9.6. – Decision support
A9.8. – Reasoning
A9.9. – Distributed AI, Multi-agent
A9.10. – Hybrid approaches for AI

Other research topics and application domains

B6.5. – Information systems
B9.5.1. – Computer science
B9.10. – Privacy

1 Team members, visitors, external collaborators

Research Scientists

- Emmanuelle Anceaume [CNRS, Senior Researcher, from Mar 2024, HDR]
- Pierre-Francois Gimenez [INRIA, ISFP, from Oct 2024]
- Yufei Han [INRIA, Senior Researcher, from Apr 2024]
- Michel Hurfin [INRIA, Researcher, HDR]
- Ludovic Mé [INRIA, Advanced Research Position, from Apr 2024, ADS, HDR]

Faculty Members

- Valerie Viet Triem Tong [Team leader, CENTRALESUPELEC, Professor, from Mar 2024, HDR]
- Christophe Bidan [CENTRALESUPELEC, Professor, from Mar 2024, HDR]
- Gilles Guette [UNIV RENNES, until Aug 2024, HDR]
- Jean-François Lalande [CENTRALESUPELEC, Professor, from Mar 2024, HDR]

Post-Doctoral Fellows

- Anatolii Khalin [CENTRALESUPELEC, Post-Doctoral Fellow, from Mar 2024]
- Omnia Mohamed [CENTRALESUPELEC, Post-Doctoral Fellow, from May 2024]
- Fabien Pesquerel [INRIA, Post-Doctoral Fellow, from Apr 2024]

PhD Students

- Lucas Aubard [INRIA, from Apr 2024]
- Pierre-Victor Besson [UNIV RENNES, ATER, from Mar 2024 until Aug 2024]
- Fanny Dijoud [INRIA, from Apr 2024]
- Sebastien Kilian [CENTRALESUPELEC, from Apr 2024]
- Maxime Lanvin [CENTRALESUPELEC, from Mar 2024]
- Pierre Lledo [DGA-MI, from Mar 2024]
- Jean-Marie Mineau [CENTRALESUPELEC, from Mar 2024]
- Matthieu Mouzaoui [INRIA, from Apr 2024]
- Hélène Orsini [CENTRALESUPELEC, from Mar 2024]
- Manuel Poisson [AMOSSYS, CIFRE, from Mar 2024]
- Vincent Raulin [CENTRALESUPELEC, ATER, from Oct 2024]
- Vincent Raulin [INRIA, from Apr 2024 until Sep 2024]
- Adrien Schoen [INRIA, from Apr 2024]
- Natan Talon [HACKUITY, CIFRE, from Mar 2024 until Nov 2024]
- Patrick Zounon [INRIA, from Oct 2024]

Technical Staff

- Yohann Rio [CENTRALESUPELEC, from Oct 2024]
- Alexandre Sanchez [INRIA]

Interns and Apprentices

- Anass Belarbi [CENTRALESUPELEC, Intern, from Mar 2024 until Oct 2024]
- Akram Blal [INRIA, Intern, from Apr 2024 until Aug 2024]
- Yohann Morel [CENTRALESUPELEC, Intern, from Mar 2024 until Aug 2024]
- David Emmanuel Ramirez Stanford [CENTRALESUPELEC, Intern, from Apr 2024 until Jul 2024]
- Gayathiri Ravendirane [CENTRALESUPELEC, Intern, from Jun 2024 until Jul 2024]
- Yohann Rio [CENTRALESUPELEC, Intern, from Mar 2024 until Aug 2024]
- Guillaume Vachez [CENTRALESUPELEC, Intern, from Jun 2024]
- Abdallah Zerkani [CENTRALESUPELEC, Intern, from Apr 2024 until Jul 2024]

Administrative Assistant

- Amandine Seigneur [INRIA, from Sep 2024]

Visiting Scientists

- Patrik Goldschmidt [UNIV TECH BRNO, from May 2024 until Aug 2024]
- Martin Mocko [CENTRALESUPELEC, from Oct 2024]

External Collaborator

- Frederic Majorczyk [DGA, from Mar 2024]

2 Overall objectives

The PIRAT team is intested in attack campaigns pursued by an attacker with the aim of perpetrating a malicious action on a specific element in a heterogeneous information system.

The team has three primary objectives in their research efforts:

- **Comprehension**: Understanding the scope and mechanisms of attacks by analyzing artifacts, logs, and malware.
- **Detection**: Developing advanced techniques to identify attacks as early as possible, using real-time observations like logs and network traffic.
- **Resistance** : Enhancing system resilience against attacks

The PIRAT team integrates expertise across programming, networks, AI, and distributed systems to propose solutions for the evolving cybersecurity landscape.

3 Research program

The research program of the PIRAT team revolves around the three main lines of research mentioned above:

Comprehension of Attacks

This line of research focuses on understanding the mechanisms and scope of cyberattacks. The objectives include:

- **Data Collection**: Acquiring up-to-date and representative attack data from sources like honeypots and attack-defense exercises.
- **Attack Modeling**: Developing tools and methods to build clear and scalable representations of complex attack scenarios. These models aim to assist experts in analyzing compromised systems, including large-scale infrastructures.

Detection of Attacks The PIRAT team aims to improve the detection of attacks by addressing current limitations. Our goals include:

- **Distributed Detection** : Designing collaborative systems that combine local analysis with global models, leveraging federated learning and fault-tolerant mechanisms to ensure privacy and robustness.
- **AI-Driven Detection**: Creating adaptable intrusion detection systems (IDS) powered by artificial intelligence, with a focus on explainable AI (XAI) and human-in-the-loop approaches to enhance operational efficiency and reduce false positives.

Resistance to Attacks Our goal is to improve the resilience of systems against attacks and ensure rapid recovery. Key objectives include:

- **Automatic Cyber Ranges**: Developing automated platforms for training and evaluating defensive tools.
- **Hardening Security Tools**: Strengthening AI-driven threat detection and providing irrefutable evidence for undesirable behaviors.
- **Containment Strategies** : Implementing mechanisms to isolate and mitigate the impact of attacks, restoring normal operations quickly.

Technological Development and Transfer The team emphasizes the development of reproducible tools and datasets:

- **Open-source publication** of mature tools and data for the academic community.
- **Industry collaborations** to create practical, transferable solutions.
- **Training programs and public engagement** to disseminate our research and tools.

4 Application domains

The PIRAT team's application domains focus on the practical use of their research findings in cybersecurity. These include: Infrastructure Protection, Intrusion Detection Systems including AI-driven, Security in AI and Distributed Systems, Malware Analysis and Mitigation, Education an Cyber training.

5 Highlights of the year

5.1 Awards

- Adrien Shoen and Pierre-Francois Gimenez have received the Best Community Contribution Award for their contribution called "FlowChronicle" presented at CoNEXT 2024, with their co-authors Joscha Cueppers and Gregory Blanc.

6 New software, platforms, open data

6.1 New software

6.1.1 TADAM

Name: TADAM

Keywords: Timed automata, Machine learning

Functional Description: Timed Automata (TA) are formal models capable of representing regular languages with timing constraints, making them well-suited for modeling systems where behavior is driven by events occurring over time. Most existing work on TA learning relies on active learning, where access to a teacher is assumed to answer membership queries and provide counterexamples. While this framework offers strong theoretical guarantees, it is impractical for many real-world applications where such a teacher is unavailable. In contrast, passive learning approaches aim to infer TA solely from sequences accepted by the target automaton. However, current methods struggle to handle noise in the data, such as symbol omissions, insertions, or permutations, which often result in excessively large and inaccurate automata. TADAM is a novel approach that leverages the Minimum Description Length (MDL) principle to balance model complexity and data fit, allowing it to distinguish between meaningful patterns and noise. TADAM is significantly more robust to noisy data than existing techniques, less prone to overfitting, and produces concise models that can be manually audited.

Release Contributions: First public version.

URL: https://github.com/Fos-R/TADAM

Contact: Pierre-Francois Gimenez

Partner: CISPA Helmholtz Center for Information Security

6.1.2 FlowChronicle

Name: FlowChronicle

Keywords: Generative AI, Data mining

Scientific Description: This software is the implementation of the method described in the article "FlowChronicle: Synthetic Network Flow Generation through Pattern Set Mining" published at the CoNEXT 2024 conference.

Functional Description: FlowChronicle creates network flows statistically similar to a network capture. It allows in particular dataset augmentation.

Release Contributions: First publicly available version.

URL: https://github.com/Fos-R/FlowChronicle

Contact: Pierre-Francois Gimenez

Partner: CISPA Helmholtz Center for Information Security

6.1.3 BAGUETTE

Keywords: Malware, Data mining

Functional Description: Malware analysis consists of studying a sample of suspicious code to understand it and producing a representation or explanation of this code that can be used by a human expert or a clustering/classification/detection tool. The analysis can be static (only the code is

studied) or dynamic (only the interaction between the code and its host during one or more executions is studied). The quality of the interpretation of a code and its later detection depends on the quality of the information contained in this representation. To date, many analyses produce voluminous reports that are difficult to handle quickly. BAGUETTE is a graph-based representation of the interactions of a sample and the resources offered by the host system during one execution. BAGUETTE helps automatically search for specific behaviors in a malware database and efficiently assists the expert in analyzing samples.

URL: https://gitlab.inria.fr/vraulin/baguette-verse

Contact: Vincent Raulin

6.1.4 URSID

Keywords: Cybersecurity, Cyber attack, Virtual Machine, Cyber Range

Functional Description: URSID makes it possible to deploy multiple variants of vulnerable virtual architectures from a single attack scenario description. These architectures can be used to train security teams or students, or as a honeypot for learning and analyzing attack techniques used in the field.

Contact: Pierre-Victor Besson

6.2 New platforms

Smart and Secure Room platform

Participants: Jean-François Lalande, Anatolii Khalin.

The Smart and Secure Room platform is one of the platforms of the LHS of Rennes. It is intended to explore attacks targeting IoT environments and energy managemetn in smart buildings. In 2024, Anatolii Khalin worked on detecting cyberattacks that could target such a physical system. In particular, smart buildings taking autonomous decisions about energy production and consumption could be the target of an attacker. We have conducted a campaign for simulating an attacker stealing energy when the smart room is producing and consuming solar energy. We propose new methods to detect such attacks and we explore attacks that may be undetected.

Poneypot platform

Participants: Yufei Han, Yohann Morel, Hélène Orsini, Gayathiri Ravendirane, Alexandre Sanchez, Guillaume Vachez.

The Poneypot platform is a multi-instance honeypot designed for collecting network and system traffic generated by Botnet activity

6.3 Open data

The team has released the following datasets and data generator:

- A dataset of static analysis tools for Android applications [10] that enables to reproduce analysis at the time of the publication of the tools.
- GothX is an open-source, replicable and highly customizable traffic generator for heterogeneous IoT networks [12].

7 New results

7.1 Axis 1 : Comprehension of Attacks

Participants: Yufei Han, Hélène Orsini, Gayathiri Ravendirane, David Emmanuel Ramirez Stanford, Guillaume Vachez, Valerie Viet Triem Tong.

Network attack attribution Network attack attribution is crucial for identifying and understanding attack campaigns, and implementing preemptive measures. Traditional machine learning approaches face challenges such as labor-intensive campaign annotation, imbalanced attack data distribution, and concept drift. To address these challenges, we propose DYNAMO [11], a novel weakly supervised and human-in-the-loop machine learning framework for au- tomated network attack attribution using raw network traffic records. DYNAMO integrates self-supervised learning and density-aware active learning techniques to reduce the overhead of exhaustive annotation, querying human analysts to label only a few selected highly representative network traffic samples. Our experiments on the CTU-13 dataset demonstrate that annotating less than 3% of the records achieves attribution accuracy comparable to fully supervised approaches with twice as many labeled records. Moreover, compared to classic active learning and semi-supervised techniques, DYNAMO achieves 20% higher attribution accuracy and nearly perfect detection accuracy for unknown botnet campaigns with minimal annotations.

Unveiling Stealthy Backdoor Attacks against Personalized Federated Learning Federated Learning (FL) is a collaborative machine learning technique where multiple clients work together with a central server to train a global model without sharing their private data. However, the distribution shift across non-IID datasets of clients poses a challenge to this one-model-fits-all method hindering the ability of the global model to effectively adapt to each client's unique local data. To echo this challenge, personalized FL (PFL) is designed to allow each client to create personalized local models tailored to their private data. While extensive research has scrutinized backdoor risks in FL, it has remained underexplored in PFL applications. In [8], we delve deep into the vulnerabilities of PFL to backdoor attacks. Our analysis showcases a tale of two cities. On the one hand, the personalization process in PFL can dilute the backdoor poisoning effects injected into the personalized local models. Furthermore, PFL systems can also deploy both server-end and client-end defense mechanisms to strengthen the barrier against backdoor attacks. On the other hand, our study shows that PFL fortified with these defense methods may offer a false sense of security. We propose PFedBA, a stealthy and effective backdoor attack strategy applicable to PFL systems. PFedBA ingeniously aligns the backdoor learning task with the main learning task of PFL by optimizing the trigger generation process. Our comprehensive experiments demonstrate the effectiveness of PFedBA in seamlessly embedding triggers into personalized local models. PFedBA yields outstanding attack performance across 10 state of-the-art PFL algorithms, defeating the existing 6 defense mechanisms. Our study sheds light on the subtle yet potent backdoor threats to PFL systems, urging the community to bolster defenses against emerging backdoor challenges.

Collusive Backdoor Attacks to Federated Learning Considerable efforts have been devoted to addressing distributed backdoor attacks in federated learning (FL) systems. While significant progress has been made in enhancing the security of FL systems, our study ([3]) reveals that there remains a false sense of security surrounding FL. We demonstrate that colluding malicious participants can effectively execute backdoor attacks during the FL training process, exhibiting high sparsity and stealthiness, which means they can evade common defense methods with only a few attack iterations. Our research highlights this vulnerability by proposing a Collusive Backdoor Attack named CoBA. CoBA is designed to enhance the sparsity and stealthiness of backdoor attacks by offering trigger tuning to facilitate learning of backdoor training data, controlling the bias of malicious local model updates, and applying the projected gradient descent technique. By conducting extensive empirical studies on 5 benchmark datasets, we make the following observations: 1) CoBA successfully circumvents 15 state-of-theart defense methods for robust FL; 2) Compared to existing backdoor attacks, CoBA consistently achieves superior attack performance; and 3) CoBA can achieve persistent poisoning effects through significantly sparse attack iterations. These

findings raise substantial concerns regarding the integrity of FL and underscore the urgent need for heightened vigilance in defending against such attacks.

Understanding Privacy Leaks in Vertical Federated Learning Systems Vertical Federated Learning (VFL) is a collaborative learning paradigm where participants share the same sample space while splitting the feature space. In VFL, local participants host their bottom models for feature extraction and collaboratively train a classifier by exchanging intermediate results with the server owning the labels. Both local training data and bottom models contain privacy-sensitive information and are considered the intellectual property of each participant, and thus should be protected by the design of VFL. In [2], we expose the fundamental susceptibility of VFL systems to privacy leaks, which arise from the collaboration between the server and clients during both training and testing. Based on our findings, we propose PISTE, a model-agnostic framework of privacy stealing attacks against VFL. PISTE delivers three privacy inference attacks, i.e., model stealing, data reconstruction, and property inference attacks on five benchmark datasets and four different model architectures. We further discuss four potential countermeasures. Experimental results show that all of them cannot prevent all three privacy stealing attacks in PISTE. In summary, our study demonstrates the inherent yet rarely uncovered vulnerability of VFL on leaking data and model privacy.

Cross-Context Backdoor Attacks against Graph Prompt Learning Graph Prompt Learning (GPL) bridges significant disparities between pretraining and downstream applications to alleviate the knowledge transfer bottleneck in real-world graph learning. While GPL offers superior effectiveness in graph knowledge transfer and computational efficiency, the security risks posed by backdoor poisoning effects embedded in pretrained models remain largely unexplored. Our study provides a comprehensive analysis of GPL's vulnerability to backdoor attacks. We introduce CrossBA, the first cross-context backdoor attack against GPL, which manipulates only the pretraining phase without requiring knowledge of downstream applications. Our investigation reveals both theoretically and empirically that tuning trigger graphs, combined with prompt transformations, can seamlessly transfer the backdoor threat from pretrained encoders to downstream applications. Through extensive experiments involving 3 representative GPL methods across 5 distinct cross-context scenarios and 5 benchmark datasets of node and graph classification tasks, we demonstrate that CrossBA consistently achieves high attack success rates while preserving the functionality of downstream applications over clean input. We also explore potential countermeasures against CrossBA and conclude that current defenses are insufficient to mitigate CrossBA. Our study highlights the persistent backdoor threats to GPL systems, raising trustworthiness concerns in the practices of GPL techniques.

7.2 Axis 2 : Detection of Attacks

Participants: Pierre-Francois Gimenez, Maxime Lanvin, Frederic Majorczyk.

Learning Timed Automata from Noisy Observations Timed Automata (TA) are formal models capable of representing regular languages with timing constraints, making them well-suited for modeling systems where behavior is driven by events occurring over time. Most existing work on TA learning relies on active learning, where access to a teacher is assumed to answer membership queries and provide counterexamples. While this framework offers strong theoretical guarantees, it is impractical for many real-world applications where such a teacher is unavailable. In contrast, passive learning approaches aim to infer TA solely from sequences accepted by the target automaton. However, current methods struggle to handle noise in the data, such as symbol omissions, insertions, or permutations, often resulting in excessively large and inaccurate automata. In [6], we introduce TADAM, a novel approach that leverages the Minimum Description Length (MDL) principle to balance model complexity and data fit, allowing it to distinguish between meaningful patterns and noise. We show that TADAM is significantly more robust to noisy data than existing techniques, less prone to overfitting, and produces concise models that can be

manually audited. We further demonstrate its practical utility through experiments on real-world tasks, such as network flow classification and anomaly detection.

FlowChronicle: Synthetic Network Flow Generation through Pattern Set Mining Network traffic datasets are regularly criticized, notably for the lack of realism and diversity in their attack or benign traffic. Generating synthetic network traffic using generative machine learning techniques is a recent area of research that could complement experimental test beds and help assess the efficiency of network security tools such as network intrusion detection systems. Most methods generating synthetic network flows disregard the temporal dependencies between them, leading to unrealistic traffic. To address this issue, we introduce FlowChronicle, a novel synthetic network flow generation tool from mined patterns and Bayesian networks. As a core component, we propose a novel pattern miner in combination with statistical models to preserve temporal dependencies. We empirically compare our method against state-of-the-art techniques on several criteria, namely realism, diversity, compliance, and novelty. This evaluation demonstrates the capability of FlowChronicle to achieve high-quality generation while significantly outperforming the other methods in preserving temporal dependencies between flows. Besides, in contrast to deep learning methods, the patterns identified by FlowChronicle are explainable, and experts can verify their soundness. Our work substantially advances synthetic network traffic generation, offering a method that enhances both the utility and trustworthiness of the generated network flows.

7.3 Axis 3 : Resistance to Attacks

Participants: Emmanuelle Anceaume, Gilles Guette, Yufei Han, Jean-François Lalande, Ludovic Mé, Jean-Marie Mineau, Manuel Poisson, Adrien Schoen, Natan Talon, Valerie Viet Triem Tong.

Automated web pentesting A wide array of techniques and tools can be employed for web application security assessment. Some methods, such as fuzzers and scanners, are partially or fully automated, offering speed and cost-effectiveness. However, these tools often fall short in detecting specific vulnerabilities like broken access control and are prone to generating false positives. On the other hand, manual processes like penetration testing, though more time-consuming and necessitating expertise, provide a more comprehensive risk assessment. To overcome the limitations of automated tools, these techniques are frequently combined. Fuzzers and scanners, despite their ease of use and quick results, require the expertise of penetration testing experts to address their limitations. By integrating these approaches, a more robust and nuanced security assessment can be achieved. In [14], we present SCWAD, an automated and customizable penetration testing framework designed to assess vulnerabilities in web applications.

Adversarial Robustness on Categorical Data Research on adversarial robustness has predominantly focused on continuous inputs, leaving categorical inputs, especially tabular attributes, less examined. To echo this challenge, our work aims to evaluate and enhance the robustness of classification over categorical attributes against adversarial perturbations through efficient attack-free approaches. In [15], we propose a robustness evaluation metric named Integrated Gradient-Smoothed Gradient (IGSG). It is designed to evaluate the attributional sensitivity of each feature and the decision boundary of the classifier, two aspects that significantly influence adversarial risk, according to our theoretical analysis. Leveraging this metric, we develop an IGSG-based regularization to reduce adversarial risk by suppressing the sensitivity of categorical attributes. We conduct extensive empirical studies over categorical datasets of various application domains. The results affirm the efficacy of both IGSG and IGSG-based regularization. Notably, IGSG-based regularization surpasses the state-of-the-art robust training methods by a margin of approximately 0.4% to 12.2% on average in terms of adversarial accuracy, especially on high-dimension datasets.

Sharding in permissionless systems in presence of an adaptive adversary Blockchain scalability is a recurrent issue. In [5, 4], we present SplitChain, a protocol intended to support the creation of scalable proof-of-stake and account-based blockchains without undermining decentralization and security. This

is achieved by using sharding, i.e. by splitting the blockchain into several lighter chains managed by their own disjoint sets of validators called shards. These shards balance the load by processing disjoint sets of transactions in parallel. SplitChain distinguishes itself from other sharded blockchains by reducing the synchronization constraints among shards while maintaining security guarantees in an asynchronous setting. A dedicated routing protocol enables transactions to be redirected between shards with a low number of hops and messages. Finally, the protocol is designed to dynamically adapt the number of shards to the system load to avoid over-dimensioning issues encountered in static sharding-based solutions.

On the formal specification of Byzantine tolerant asset transfert . We are currently studying a new asynchronous Byzantine-tolerant asset transfer system (cryptocurrency) with three noteworthy properties: quasi-anonymity, lightness, and consensus-freedom in presence of Byzantine adversaries [18]. Quasi-anonymity means no information is leaked regarding the receivers and amounts of the asset transfers. Lightness means that the underlying cryptographic schemes are *succinct*, and each process only stores data polylogarithmic in the number of its own transfers. Consensus-freedom means the system does not rely on a total order of asset transfers. The proposed algorithm is the first asset transfer system that simultaneously fulfills all these properties in the presence of asynchrony and Byzantine processes. To obtain them, the paper adopts a modular approach combining a new distributed object called agreement proofs and well-known techniques such as vector commitments, universal accumulators, and zero-knowledge proofs. The paper also presents a new non-trivial universal accumulator implementation that does not need knowledge of the underlying accumulated set to generate (non-)membership proofs, which could benefit other crypto-based applications.

Compressing permissionless blockchains The long-term feasibility of blockchain technology is hindered by the inability of existing blockchain protocols to prune the consensus data leading to constantly growing storage and communication requirements. Kiayias et al. have proposed Non-Interactive-Proofs-of-Proof-of-Works (NIPoPoWs) as a mechanism to reduce the storage and communication complexity of blockchains to O(poly log(n)). However NiPoPoWs are only proven to operate securely in a setting with constant difficulty. Tackling the problem of blockchain compression operating in a O(poly log(n)) storage complexity and O(poly log(n)) communication complexity with variable difficulty is a challenging problem. In [19] we present for the first time a construction that provably satisfies the requirements of a non-interactive proof of proof-of-work that is secure against a 1/3 adversary in a dynamic environment. Succinctness, security and onlineness of the scheme are proven, while experimental results confirm the exponential reduction of the Bitcoin blockchain.

7.4 Reproducibility, reusability and open data

Participants: Pierre-Francois Gimenez, Maxime Lanvin, Jean-François Lalande, Jean-Marie Mineau, Ludovic Mé, Manuel Poisson, Adrien Schoen.

Reproducibility and reusability in computer science experiments become a requirement for research works. Reproducibility ensures that results can be confirmed by using the same dataset and software of previous papers. Reusability helps other researchers to build new approaches with distributed software artifacts.

Evaluating the Reusability of Android Static Analysis Tools For researchers in the field of security of mobile platforms, ensuring reproducibility and reusability is difficult to implement. In particular for reusability, datasets of Android applications may contain recent applications that past analysis software cannot process. As a consequence, past software produced by researchers may be difficult to reuse, which endangers the reproducibility of research. We have explored the reusability of past software dedicated to static analysis of Android applications. We extensively evaluated the success or failure of these past software on a dataset containing Android applications that can have up to six years of distance from the

original publication. We also measured the influence of some important characteristics of the application such as being a goodware or a malware or the application size.

Customizable, legitimate and malicious IoT network traffic generation In recent years, machine learning-based anomaly detection (AD) has become an important measure against security threats from Internet of Things (IoT) networks. Machine learning (ML) models for network traffic AD require datasets to be trained, evaluated and compared. Due to the necessity of realistic and up-to-date representation of IoT security threats, new datasets need to be constantly generated to train relevant AD models. Since most traffic generation setups are developed considering only the author's use, replication of traffic generation becomes an additional challenge to the creation and maintenance of useful datasets. In [12], we propose GothX, a flexible traffic generator to create both legitimate and malicious traffic for IoT datasets. As a fork of Gotham Testbed, GothX is developed with five requirements: 1)easy configuration of network topology, 2) customization of traffic parameters, 3) automatic execution of legitimate and attack scenarios, 4) IoT network heterogeneity (the current iteration supports MQTT, Kafka and SINETStream services), and 5) automatic labeling of generated datasets. GothX is validated by two use cases: a) re-generation and enrichment of traffic from the IoT dataset MQTTset,and b) automatic execution of a new realistic scenario including the exploitation of a CVE specific to the Kafka-MQTT network topology and leading to a DDoS attack. We also contribute with two datasets containing mixed traffic, one made from the enriched MQTTset traffic and another from the attack scenario. We evaluated the scalability of GothX (450 IoT sensors in a single machine), the replication of the use cases and the validity of the generated datasets, confirming the ability of GothX to improve the current state-of-the-art of network traffic generation.

Synthetic network traffic generation The evaluation of network intrusion detection systems requires a sufficient amount of mixed network traffic, i.e., composed of both malicious and legitimate flows. In particular, obtaining realistic legitimate traffic is hard. Synthetic network traffic is one of the tools to respond to insufficient or incomplete real-world datasets. In [13], we only focus on synthetically generating high-quality legitimate traffic and we do not delve into malicious traffic generation. For this specific task, recent contributions make use of advanced machine learning-driven approaches, notably through Generative Adversarial Networks (GANs). However, evaluations of GAN-generated data often disregards pivotal attributes, such as protocol adherence. Our study addresses the gap by proposing a comprehensive set of metrics that assess the quality of synthetic legitimate network traffic. To illustrate the value of these metrics, we empirically compare advanced network-oriented GANs with a simple and yet effective probabilistic generative model, Bayesian Networks (BN). According to our proposed evaluation metrics, BNbased network traffic generation outperforms the state-ofthe-art GAN-based opponents. In our study, BN yields substantially more realistic and useful synthetic benign traffic and minimizes the computational costs simultaneously.

8 Bilateral contracts and grants with industry

8.1 Bilateral contracts with industry

DGA

Participants: Pierre-Francois Gimenez, Yufei Han, Vincent Raulin, Valerie Viet Triem Tong, Alexandre Sanchez.

Vincent Raulin's PhD focuses on using machine learning approaches to boost malware detection and classification based on dynamic analysis traces by extracting feature representations with the knowledge of malware analysis experts. This representation aims at capturing the semantics of the program (i.e., what resources it accesses, what operations it performs on them) in a plateform-independent fashion, by replacing the implementation particularities (system call number 2) with higher-level operation (opening a file). This representation could notably provide semantic explanation of malware activity and deliver explainable malware detection/malware family classification.

DGA

Participants: Jean-François Lalande, Pierre Lledo, Frederic Majorczyk.

Pierre Lledo has started his phD thesis on the evaluation of simulators of attacks in emulated information systems.

AMOSSYS

Participants: Manuel Poisson, Gilles Guette, Valerie Viet Triem Tong.

Manuel Poisson has started a thesis in collaboration with Amossys. Manuel Poisson is interested in identifying operational attack scenarios in an information system.

Hackuity

Participants: Natan Talon, Gilles Guette, Yufei Han, Valerie Viet Triem Tong.

Natan Talon started his PhD in October 2021 in the context of a collaboration with the company Hackuity. The main objective of this thesis is to be able to assess whether an information system is likely to be vulnerable to an attack. This attack may have been observed in the past or inferred automatically from other attacks.

8.2 Bilateral Grants with Industry

DGA

Participants: Fanny Dijoud, Pierre-Francois Gimenez, Michel Hurfin, Frederic Majorczyk.

The objective of Fanny Dijoud's thesis is to design an intrusion detection system to analyze system logs. The approach is based on the use of several AI models to observe anomalies in the different provenance graphs that are built. The proposed solution takes into account the fact that these graphs are dynamic and heterogeneous. This PhD has started in november 2023.

ANSSI

Participants: Lucas Aubard, Gilles Guette, Ludovic Me.

Lucas Aubard started his PhD in October 2022 in the context of a collaboration between Inria and the ANSSI. The objective of this thesis is to improve the existing knowledge on reassembly policies, to design mechanisms to automate IDS configuration and to improve the application of these policies within IDS/IPS to increase their detection capabilities in specific contexts such as cloud computing.

DGA

Participants: Pierre-Francois Gimenez, Maxime Lanvin, Frederic Majorczyk, Ludovic Me.

Maxime Lanvin is financed by the DGA through the Pôle d'Excellence Cyber (PEC) since October 2021. Maxime works on behavorial intrusion detection based on machine learning techniques. His work focuses on the analysis of time series to detect APT attacks.

DGA

Participants: Pierre-Francois Gimenez, Yufei Han, Frederic Majorczyk, Ludovic Me, Adrien Schoen.

Adrien Schoen is financed by the DGA though the Pôle d'Excellence Cyber (PEC) since October 2021. Adrien works on the generation of synthetic network dataset to better evaluate intrusion detection systems. This work is based on various deep learning models such as generative adversarial network and variational auto-encoder.

DGA

Participants: Yufei Han, Valerie Viet Triem Tong.

Helene Orsini's PdD thesis is financed by DGA since October 2021. Her thesis project focuses on adversarially robust and interpretable machine learring pipeline for network intrusion detection systems. She will study how to automate the feature engineering phase to extract informative features from non-structured, categorical and imperfect security reports / logs. Furthermore, she will investigate how to make the machine learning pipeline resilient to intentional evading techniques in network intrusion behaviors.

9 Partnerships and cooperations

9.1 International initiatives

9.1.1 Inria associate team not involved in an IIL or an international program

SecGen

Title: Security Data Generation

Duration: 2023 -> 2025

Coordinator: Mario Fritz (fritz@cispa.de)

Partners:

- CISPA (Allemagne)

Inria contact: Pierre-francois Gimenez

Summary: SecGen aims at improving network traffic generation to make it usable for intrusion detection systems learning and evaluation. Indeed, existing datasets are highly criticized because of their age, their non-representativeness and the errors they contain. The PIRAT team is already involved in a thesis on network data generation, and CISPA will bring their expertise in data mining (for the identification of temporal patterns specific to network data that our experts will be able to analyze) and in anomaly detection based on deep learning.

9.2 International research visitors

9.2.1 Visits of international scientists

Other international visits to the team
Lénaïg Cornanguer

Participants: Pierre-Francois Gimenez.

Status Post-doc

Institution of origin: CISPA

Country: Germany

Dates: October 2024 (1 week)

Context of the visit: identification of new research questions in the context of the SecGen associate team

Mobility program/type of mobility: research stay

Patrik Goldschmidt

Participants: Pierre-Francois Gimenez, Yufei Han.

Status PhD

Institution of origin: KInIT

Country: Slovaquia

Dates: May 2024 to August 2024 (3 months)

Context of the visit: Joint work on transfer learning applied to network intrusion detection

Mobility program/type of mobility: research stay

Martin Mocko

Participants: Yufei Han, Martin Mocko, Vincent Raulin, Alexandre Sanchez, Valerie Viet Triem Tong.

Status PhD

Institution of origin: Brno University of Technology

Country: Slovaquia

Dates: November 2024 to December 2024 (2 months)

Context of the visit: Dynamic analysis of malware. Collaboration in the context of the PEPR Defmal project.

Mobility program/type of mobility: research stay

9.3 National initiatives

PEPR CyberSecurity project: DefMal (2022-2028)

Participants: Jean-François Lalande, Yufei Han, Valerie Viet Triem Tong.

PEPR DefMal is a collaborative ANR project involving CentraleSupélec, Rennes University, Lorraine University, Sorbonne Paris Nord University, CEA, CNRS, Inria and Eurecom. Malware is affecting government systems, critical infrastructures, businesses, and citizens alike, and regularly makes headlines in the press. Malware extorts money (ransomware), steals data (banking, medical), destroys information systems, or disrupts the operation of industrial systems. The fight against malware is a national and European security issue that requires scientific advances to design new responses and anticipate future attack methods. The aim of the project DefMal is to study malicious programs, whether they are malware, ransomware, botnet, etc. The first objective is to develop new approaches to analyze malicious programs. This objective covers the three aspects of the fight against malware: (i) Understanding (ii) Detection and (iii) Forensics. The second objective of the project is the global understanding of the malware ecosystem (modes of organization, diffusion, etc.) in an interdisciplinary approach involving all the actors concerned.

PEPR CyberSecurity project: SuperViz (2022-2028)

Participants: Pierre-Francois Gimenez, Gilles Guette, Yufei Han, Maxime Lanvin, Frederic Majorczyk, Ludovic Mé, Yohann Morel, Adrien Schoen.

PEPR SuperviZ is a collaborative ANR project involving CentraleSupélec, Eurecom, Institut Mines-Télécom, Institut Polytechnique de Grenoble, Rennes University, Lorraine University, CEA, CNRS and Inria. The digitalization of all infrastructures makes it almost impossible today to secure all systems a priori, as it is too complex and too expensive. Supervision seeks to reinforce preventive security mechanisms and to compensate for their inadequacies. Supervision is fundamental in the general context of enterprise systems and networks, and is just as important for the security of cyber-physical systems. Indeed, with the ever growing number of connections between objects, the attack surface of systems has become frighteningly wide. This makes security even more difficult to implement. The increase in the number of components to be monitored, as well as the growing heterogeneity of the capacity of these objects in terms of communication, storage and computation, makes security supervision more complex.

PEPR CyberSecurity project: Rev (2023-2028)

Participants: Pierre-Francois Gimenez.

PEPR REV is a project about vulnerability research and exploitation. A notable characteristic of complex targets is that they can generally no longer be attacked using a single technique or exploiting a single vulnerability, due to the deployment of numerous protections. For this reason, the REV project is tackling this problematic at multiple levels by addressing all layers, hardware, software and communication interfaces (web and IoT). In this purpose, one of the project's objectives is to combine several tools and approaches simultaneously: for example, memory analysis will benefit from advances in hardware attacks, and will be used to develop exploits. This broad-spectrum analysis is fundamental today: as an illustration, hardware attacks can be combined with software attacks, software attacks can be based on weaknesses in the micro-architecture or require advanced network interactions. Moreover, the impact of attacks and exploits nowadays goes far beyond malicious use, allowing for instance to forensically investigate complex systems such as smartphones. The question also arises from an ethical and legal point of view, and this is a major societal issue: to which extent is it possible to use these techniques, in particular for law enforcement, from an ethical or legal point of view. What is the possible use of these attacks, when should they be corrected ("responsible disclosure") or used, and in what legal framework?

ANR project: CKRISP (2024-2027)

Participants: Yufei Han, Michel Hurfin, Frederic Majorczyk, Patrick Zounon
.

CKRISP is an collaborative project led by Yufei Han in PIRAT with Eurecom, CEA, Telecom Sud-Paris. The main contribution of CKRISP address the limited coverage of attack behaviour variety in the training data as well as the lack of interpretability of AI detection models. First, we will investigate the combination of AI systems such as, e.g., Large Language Models (LLM), and human-monitored cyber security knowledge graph (CSKG) for understanding, predicting and exploring new cyberattack behaviours via Human-AI interaction. The powerful LLMs can help identify entities and predict relations between entities from cyber threat reports and low-level run-time behaviour logs. CSKG can then be built automatically based on extracted knowledge about specific attack scenarios. The attack knowledge graph can substantially help human analysts verify the AI-based attack detection results and facilitate human analysts' inspection of new attacks. Second, we will elaborate further on the prediction of attack behaviours by organizing AI-assisted reasoning with inputs from security incidents collected from various sensors, like IDS, and manual inspection results of human analysts. It will help assess the vulnerability of a target IT system and reach an initial step of AI-assisted security response based on the detected incidents. Third, we will further propose data generation methods to produce synthetic normal/attack behaviour data to enrich training data and improve the robustness of AI-based detection methods based on the extracted knowledge representation and causalities of attacks. Finally, new visualisation and interaction interfaces will be developed in this project to simplify the human-AI interaction. These interfaces are expected to be intuitive and user-friendly so that both technical staff (analysts) and non-technical staff (managers) can effectively interact with the results, respond quickly, and make more efficient decisions.

ANR project: Byblos(2021-2025)

Participants: Emmanuelle Anceaume.

Byblos is a collaborative ANR project involving Rennes university and IRISA (CIDRE (now PIRAT) and WIDE research teams), Nantes university (GDD research team), and Insa Lyon, LIRIS (DRIM research team). This project aims at overcoming performance and scalability issues of blockchains, that are inherent to the total order that blockchain algorithms seek to achieve in their operations, which implies in turn a Byzantine-tolerant agreement. To overcome these limitations, this project aims at taking a step aside, and exploiting the fact that many applications – including cryptocurrencies – do not require

full Byzantine agreement, and can be implemented with much lighter, and hence more scalable and efficient, guarantees. This project further argues that these novel Byzantine-tolerant applications have the potential to power large-scale multi-user online systems, and that in addition to Byzantine Fault Tolerance, these systems should also provide strong privacy protection mechanisms, that are designed from the ground up to exploit implicit synergies with Byzantine mechanisms.

ANR Project: BC4SSi (2023-2027)

Participants: Emmanuelle Anceaume.

BC4SSI is a JCJC ANR project led by Romaric Ludinard (SOTERN), involving the SOTERN and CIDRE (now PIRAT) research teams. Self-sovereign identities (SSI) are digital identities that are managed in a decentralized manner. This technology allows users to self-manage their digital identities without depending on third-party providers to store and centrally manage the data, including the creation of new identities. Implementing SSI requires a lot of care since identities are more than simple identifiers: they need to be checked by the service provider via, for instance, verifiable claims. Such requirements make blockchain technology a prime candidate for deploying SSI and storing verifiable claims. BC4SSI aims at studying the weakest synchrony assumptions enabling SSI deployment in a public Blockchain. Among the different existing challenges, BC4SSI will address the following scientific locks: alternatives to PoW security proofs, lightweight replication, scalability and energy consumption.

CMA project : Train-Cyber-Expert (TCE) (2022-2026)

Participants: Yohann Rio, Alexandre Sanchez, Valerie Viet Triem Tong.

As part of the France 2030 recovery plan, we are involved in the Train-Cyber-Expert (TCE) project, funded by the CMA (Competences and Jobs of the Future) call for projects. TCE is a collaborative project involving several academic partners with the aim of developing educational resources in the form of digital content and technological platforms, organized into skill blocks, with a focus on modularity, reusability, and competency-based pedagogy leading to certifications. We are involved in this project together with the INRIA SUSHI team. Our goal is to propose pedagogical ressources in the field of system security.

BPI project : SECURITY TWIN (2024-2027)

Participants: Manuel Poisson, Gilles Guette, Valerie Viet Triem Tong.

The Security Twin project, developed in collaboration with Amossys, a company specializing in cybersecurity, aims to enhance the security of information systems through the creation of a digital twin. This digital twin will faithfully replicate a real information system (IS), incorporating its configurations and vulnerabilities, to simulate realistic attack scenarios and assess their impact.

This project involves several challenges, both technical and scientific, such as:

- Modeling a security digital twin.
- Automating its deployment.
- Developing an attack agent capable of testing the resilience of IS under real-world conditions.

9.4 Regional initiatives

Smart and Secure Room project

Participants: Anatolii Khalin, Jean-François Lalande.

Anatolii Khalin started in November as a post-doctoral researcher in the team, co-supervized with the AUT team from IETR. His work focuses on detecting cyberattacks that could target a cyberphysical system. In particular, smart buildings taking autonomous decisions about energy production and consumption could be the target of an attacker. We plan to design new estimators used to predict the different physical measures of a smart building. These estimators could be used to raise alerts when a deviation from the expected prediction is detected, for example, because of a compromised device in the building.

CominLabs project: Priceless (2021-2025)

Participants: Emmanuelle Anceaume.

Priceless is a collaborative CominLabs project involving Rennes University with IRISA (CIDRE and WIDE research teams), and IODE (Institut de l'ouest: droit et Europe), and Nantes university (GDD research team). Promoters of blockchain-based systems such as cryptocurrencies have often advocated for the anonymity these provide as a pledge of privacy protection, and blockchains have consequently been envisioned as a way to safely and securely store data. Unfortunately, the decentralized, fully-replicated and unalterable nature of the blockchain clashes with both French and European legal requirements on the storage of personal data, on several aspects such as the right of rectification and the preservation of consent. This project aims to establish a cross-disciplinary partnership between Computer Science and Law researchers to understand and address the legal and technical challenges associated with data storage in a blockchain context.

DECOR project: (2023-2025)

Participants: Jean-François Lalande, Omnia Mohamed.

PIRAT has participated to the I-Demo project DECOR with Wallix and Malizen. The goal of the project is to propose automatic mitigations when an attack occurs in an information system managed by SCAR, a tool for managing heterogenous firewalls. After detecting an security incident, an analyst would investigate the collected log using the Malizen platform. Based on these logs and the analysts actions, we intend to propose automatic mitigations, for example to contain the attack in a subpart of the subnet of the information system. The first year of the project has been dedicated to the collect of the logs and the rebuild of the topology of the infrastructure.

10 Dissemination

Participants: Emmanuelle Anceaume, Christophe Bidan, Pierre-Francois Gimenez, Gilles Guette, Yufei Han, Michel Hurfin, Jean-François Lalande, Frederic Majorczyk, Ludovic Mé, Hélène Orsini, Valerie Viet Triem Tong.

10.1 Promoting scientific activities

10.1.1 Scientific events: organisation

General chair, scientific chair Emmanuelle Anceaume has been co-general chair of the 7th Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS 2024), Berlin.

Member of the organizing committee Ludovic Mé served in the steering committee of RESSI (Rendez-Vous de la Recherche et de l'Enseignement de la Sécurité des Systèmes d'Information).

10.1.2 Scientific events: selection

Chair of conference program committees Valérie Viet Triem Tong has been co-chair of the program committee for the 17th International Symposium on Foundations & Practice of Security (FPS – 2024)

Hélène Orsini has been co-chair of the program committee for the Phd-thesis session of RESSI (Rendez-Vous de la Recherche et de l'Enseignement de la Sécurité des Systèmes d'Information).

Gilles Guette served as Scientific Track Chair of International Conference on Information and Communication Technologies for Disaster Management (ICT-DM 2024).

Member of the conference program committees Jean-Francois Lalande has been part of the technical program committee of:

- EICC 2024: European Interdisciplinary Cybersecurity Conference
- SSTIC 2024: Symposium sur la sécurité des technologies de l'information et des communications
- IWCC 2024: 13th International Workshop on Cyber Crime
- CUING 2024: 8th International Workshop on Cyber Use of Information Hiding

Emmanuelle Anceaume has been part of the technical program committee of:

- OPODIS 2024: 28th International Conference on Principles of Distributed Systems
- SSS 2024: 26th International Symposium on Stabilization, Safety, and Security of Distributed Systems
- BRAINS 2024: 7th Conference on Blockchain Research & Applications for Innovative Networks and Services
- NCA 2024: 22nd International Symposium on Network Computing and Applications

Gilles Guette has been part of the technical program committee of ICISSP 2025 11th International Conference on Information Systems Security and Privacy

Yufei Han has been part of the technical program committee and senior program committee of:

- IJCAI 2024: Inentional Joint Conference on Artificial Intelligence
- AAAI 2024: AAAI Conference on Artificial Intelligence
- ICML 2024: International Conference on Machine Learning
- NeurIPS 2024: Conference on Neural Information Processing Systems
- ICLR 2025: International Conference on Learning representations
- CCS 2025: The ACM Conference on Computer and Communications Security
- ICDE 2025: IEEE International Conference on Data Engineering (ICDE)
- PAKDD 2025: The Pacific-Asia Conference on Knowledge Discovery and Data Mining
- SDM 2025: SIAM International Conference on Data Mining
- IJCAI 2025: Inentional Joint Conference on Artificial Intelligence

Reviewer Valérie Viet Triem Tong has served as reviewer for FPS 2024 and SECITC 2024.

Michel Hurfin has served as reviewer for CARI 2024.

Yufei Han has served as a reviewer for IEEE Transactions of Information Security (TIFS), IEEE Transactions of Dependable and Secured Computing (IEEE TDSC), IEEE Transactions of Neural Networks and Learning Systems (TNNLS), ACM Transactions on Privacy and Security (TOPS) and Comuter Security (Elsevier)

Pierre-François Gimenez served as a reviewer for ARTMAN 2024, ERTS 2024, THCon 2024.

10.1.3 Journal

Reviewer - reviewing activities Jean-Francois Lalande has been a reviewer for:

- IEEE Transactions on Reliability
- ACM Transactions on Software Engineering and Methodology
- Journal of Information Security and Applications, Elsevier
- International Journal of Information Security, Springer Nature
- Transactions on Information Forensics & Security, IEEE Computer Society

Emmanuelle Anceaume was reviewer for

- Journal of Parallel and Distributed Computing
- IEEE Transactions on Parallel and Distributed Systems
- Journal of Computers
- Theoretical Computer Science Journal

Gilles Guette has been a reviewer for Journal of Networks
Pierre-François Gimenez was reviewer for:

- Computers and Security
- Journal of Computer Security

10.1.4 Scientific expertise

Jean-Francois Lalande was a reviewer for the PhD grants of Normandie University.

Emmanuelle Anceaume co-authored an opinion for the Académie des Technologies. La Blockchain : une technologie disruptive avec des enjeux de sûreté, résilience et impact environnemental Avis de l'Académie des Technologies. co-writters are S. Abiteboul, J.-C. André and Olivier Pironneau from the Académie des sciences, A. Benveniste, Ch. de Boissieu, Th. Chevalier, M. Laroche, G. Roucairol and S. Proust from the Académie des technologies, and D. Augot DR Inria.

10.1.5 Research administration

- Jean-François Lalande was member of the recruitment committees for an Assistant Professor and a Professor position at CentraleSupelec.
- Jean-François Lalande was member of the recruitment committees for a Professor position at university of Rennes.
- Valérie Viet Triem Tong was member of the recruitment committees for an Assistant Professor and a Professor position at CentraleSupelec.
- Valérie Viet Triem Tong was member of the recruitment committees for an Assistant Professor position at Insa Toulouse.

- Valérie Viet Triem Tong was member of the recruitment committee for an Assistant Professor position at Telecom Paris.
- Valérie Viet Triem Tong was member of the recruitment committees for a Professor position at University of Brest.
- Valérie Viet Triem Tong was member of the recruitment committees for a Professor position at INP Grenoble.
- Emmanuelle Anceaume was member of the recruitment committee for an Assitant Professor position at INP Grenoble.
- Emmanuelle Anceaume was member of the recruitment committee for an Assitant Professor position at Sorbonne Université.

10.2 Teaching - Supervision - Juries

Yufei Han has given training courses on "Machine Learning and Cybersecurity: Why and How with scikit-learn" via Inria Academy. machine-learning-and-cybersecurity-why-and-how-with-scikit-learn/

10.2.1 Teaching

Several team members are involved in initial and continuing education in CentraleSupélec, a French institute of research and higher education in engineering and science, ESIR (Ecole Supérieure d'Ingénieur de Rennes) the graduate engineering school of the University of Rennes, and Centrale Casablanca.

In these institutions,

- Christophe Bidan is the head of the teaching program of Centrale Casablanca;
- Jean-François Lalande is responsible of the major program dedicated to information systems security and the special track Infosec of CentraleSupélec engineering education; He is also involved in the organization committee of EUR CyberSchool and in the computer science master degree (SIF and Cyber tracks).
- Valérie Viet Triem Tong is responsible of the mastère spécialisé (post-graduate specialization degree co-dellivered with IMT-Atlantique) in Cybersecurity for CentraleSupelec. This education was awarded best French master degree in the category "Master Cybersecurity masters and Security of systems" in the Eduniversal master ranking 2023.
- Gilles Guette was co responsible of the Systèmes Numériques et Réseaux engineer track of ESIR till august 2024.

10.2.2 Supervision

Phd defended:

- Pierre-Victor Besson, CHOUCHEN : Complete HOneynet with User Copycat on Hypervisor with Emulated Network, started November 2020, supervised by Valérie Viet Triem Tong (25%), Gilles Guette (25%), Guillaume Piolle (25%) and Erwan Abgrall (25%).
- Romain Brisse, Exploration recommendations for the investigation of security incidents, started december 2020, co-supervised by Jean-Francois Lalande (33%), Frédéric Majorczyk (33%), and Simon Boche (33%).
- Maxime Lanvin, Advancing Network Intrusion Detection: Datasets, Testbeds, Alert Explanations, and False Positives Reduction, started October 2021, supervised by Christophe Bidan (25%), Ludovic Mé (25%), Pierre-François Gimenez (25%), and Eric Totel (25%).
- Arthur Rauch, Towards more scalable and privacy-preserving distributed asset-transfer systems, started October 2021, supervised by Emmanuelle Anceaume (50%) and Davide Frey (50%).

- Adrien Schoen, Generation of realistic activities for Intrusion Detection Systems evaluation, started October 2021, supervised by Ludovic Mé (25%), Gregory Blanc (25%), Yufei Han (25%), and Frédéric Majorczyk (25%).

PhD in progress:

- Pierre Lledo, On intrusion detection, started December 2023, supervised by Jean-François Lalande (50%) and Frederic Majorczyk (50%).
- Lucas Aubard, Ambiguïtés de recouvrement de données dans les protocoles d'Internet et supervision reseau, started October 2022, supervised by Pierre Chifflier (25%), Gilles Guette (25%), Johan Mazel (25%) and Ludovic Mé (25%).
- Fanny Dijoud, Détection d'intrusions au niveau système d'informations : détection d'anomalies par traitement IA dans des graphes dynamiques hétérogènes représentant l'activité du système, started november 2023, supervised by Michel Hurfin (25%), Pierre-François Gimenez (25%), Frédéric Majorczyk (25%) et Barbara Pilastre (25%, DGA).
- Sébastien Kilian, From offensive data collection to automatic attack agent, started February 2024, supervised by Jean-François Lalande (50%), Valérie Viet Triem Tong (50%).
- Chaoran Li, Secure Artificial Intelligence for the Smart Grid Energy Management, started November 2024, supervised by Anne Blavette (50%, IETR), Michel Hurfin (25%), and Yufei Han (25%).
- Jean-Marie Mineau, Android Malware Manipulation for Improved Investigations, started November 2022, supervised by Jean-Francois Lalande (75%), Valérie Viet Triem Tong (25%).
- Yohann Morel, Adaptative Honeypot, started in October 2024, supervised by Gilles Guette (50%) (now professor at IMT) and Valérie Viet Triem Tong (50%).
- Matthieu Mouzaoui, Adversarially Robust Machine Learning-based Network Intrusion Detection System, started February 2024, supervised by Yufei Han (50%), Michel Hurfin (25%), and Gabriel Rilling (25%, CEA).
- Hélène Orsini, IA based supervision, started October 2021, supervised by Yufei Han (50%) Valérie Viet Triem Tong (25%), David Lubicz (25%).
- Dorian Pacaud, Towards blockchain frugality, started October 2024, supervised by Emmanuelle Anceaume.
- Manuel Poisson, Évaluation automatisée du niveau de sécurité d'un système d'information, started March 2023, supervised by Valérie Viet Triem Tong (25%), Gilles Guette (25%), Frédéric Guihéry (25%) and Damien Crémilleux (25%).
- Vincent Raulin, IA-based classification of malware, started October 2021, supervised by Valérie Viet Triem Tong (25%), Yufei Han (25%), Pierre-François Gimenez (50%).
- Natan Talon, Rejeu et apprentissage de scénarios d'attaques, started December 2021, supervised by Mathieu Jaume (25%), Gilles Guette (25%), Yufei Han (25%) and Valérie Viet Triem Tong (25%).
- Grégor Quetel, Détection d'anomalie et création d'une sonde d'inférence sémantique, started Octobre 2023, supervised by Pierre-François Gimenez (25%), Eric Alata (25%), Thomas Robert (25%) and Laurent Pautet (25%).
- Patrick Zounon, Constructing Cyber Security Knowledge Encoding and Reasoning from Heterogeneous Sources with Large Language Models, started October 2024, supervised by Yufei Han (50%), Michel Hurfin (25%), and Frédéric Majorczyk (25%, DGA).

10.2.3 Juries

Emmanuelle Anceaume was reviewer of Célia Mahamdi's PhD thesis. Her PhD thesis is from the Sorbonne Université, and is entitled "Multi-Consensus distribué : agrégation et révocabilité". Her defense took place on the 2nd of December 2024. It was advised by Jonathan Lejeune (MdC Sorbonne Université), Pierre Sens (Professor Sorbonne Université), Julien Sopena (MdC Sorbonne Université) and Mesaac Makpangou (CR Inria). Jury Emmanuelle Anceaume (DR CNRS, IRISA), Gil Utard (Professor Université de Picardie Jules Verne), Frédéric Le Mouël (Professor INSA Lyon), Jonathan Lejeune (MdC Sorbonne Université), Pierre Sens (Professor Sorbonne Université), Julien Sopena (MdC Sorbonne Université) et Mesaac Makpangou (CR Inria).

Emmanuelle Anceaume was reviewer of Wassim Yahyaoui's PhD thesis from the University of Luxembourg. His PhD thesis is entitled " Fast, Fair, and Secure Hierarchical Consensus and Dissemination for Blockchain Resilience despite Data-Center Disasters", and was supervised by Marcus Volp (Associate professor at Université du Luxembourg). His defense took place on the 9th of January, 2025. Jury: Alireza Esfahani (Senior Lecturer University West London), Jérémy Decouchant (Assistant Professor Delft University of Technology), Cristina Nita-Rotaru (Professeur at Northeastern University), Gabriele Lenzini (Associate professor Université du Luxembourg), Gilbert Fridgen (Professeur Université du Luxembourg).

Emmanuelle Anceaume was president of Wédan Emmanuel GNIBGA's PhD thesis defense. He received his PhD thesis from the Université de Rennes. His thesis was entitled *Modeling and optimization of Edge infrastructures and their electrical systems* . He was supervised by Anne-Cécile Orgerie (DR CNRS, IRISA) and Anne Blavette (CR CNRS). His defense took place on the 19th of November 2024. Jury: Pascal Felber (Professor Université de Neuchâtel), Pierre Sens (Professor Sorbonne Université), Emmanuelle Anceaume (DR CNRS, IRISA), Hamid Gualous (Professor Université de Caen), Anne-Cécile Orgerie (DR CNRS, IRISA) et Anne Blavette (CR CNRS).

Emmanuelle Anceaume was examinator of Joao Paulo Bezera de Araujo's PhD thesis. His thesis is from the Institut Polytechnique de Paris, and is entitled *Scalable and Reliable Decentralized Computing: From Asset Transfer to Atomic Snapshot.* He was supervised by Petr Kuznetsov (Professor Télécom Paris). His PhD thesis was defended on the 14th of November 2024. His thesis was Jury: Gaël Thomas (DR Inria Saclay), Alysson Neves Bessani (Assistant professor Universidade de Lisboa), Hagit Attiya (Professor Technion), Emmanuelle Anceaume (DR CNRS Irisa), Constantin Enea (Professor Ecole Polytechnique), Petr Kuznetsov (Professor Télécom Paris).

Emmanuelle Anceaume was examinator of Avinandan DAS's PhD thesis. His thesis is from the Université Paris Cité and is entitled *Computation with Partial Information, An Algorithmic Investigation of Graph Problems in Distributed Computing and Streaming.* He was supervised by Pierre Fraigniaud (DR CNRS IRIF) and Adi Rosen (DR CNRS, IRIF). His PhD thesis was defended on the 16th of December 2024. Jury: Emmanuelle Anceaume (DR CNRS Irisa), Marthe Bonami (CR CNRS Labri), Shay Kutten (Professeur Technion Israel), Pierre Fraigniaud (DR CNRS IRIF), Adi Rosen (DR CNRS, IRIF).

Jean-Francois Lalande has been the president of the jury for the phD defense of François Bonnal the 17th of April 2024, Mines de Saint-Etienne, untitled “Protection logicielle contre les attaques en fautes visant les microcontrôleurs utilisés dans l'internet des objets”.

Ludovic Mé was member of the PhD committee for the following PhD thesis:

- Estelle Hotellier, *Specification-based Intrusion Detection for hybrid Industrial control systems*, Université Grenoble Alpes.
- Marwan Abbas Escribano, *Modélisation de systèmes de leurres complexes*, Institut Polytechnique de Paris.
- Quentin Ducasse, *Sécurisation matérielle de la compilation à la volée des machines virtuelles langage*, ENSTA Bretagne.

Valérie Viet Triem Tong was

- reviewer of Estelle Hotellier PhD thesis, *Specification-based Intrusion Detection for hybrid Industrial control systems*, Université Grenoble Alpes.
- reviewer of Arie Haenel PhD thesis, *Hybrid security solutions for IoT devices*, Telecom Paris.

- examinator of Thomas Vigouroux Phd thesis, *Quantitative analysis for adaptive attackers*, Université Grenoble Alpes.

Pierre-François Gimenez was examinator of the following thesis:

- Hamdi Friji, "Détection d'intrusion basée sur les Graph Neural Networks pour la sécurisation des réseaux en périphérie", Télécom SudParis
- Léo Lavaur, "Améliorer la détection d'intrusions dans les systèmes répartis grâce à l'apprentissage fédéré", IMT Atlantique

Yufei Han was the examinator of the PhD thesis of Robin Duraz, *Trustable Machine Learning for Intrusion Detection Systems*, IMT Atlantique.

10.3 Popularization

Yufei Han gave a talk on *IA-based malware analysis*, at the CyberSchool Research School in 2024.

Valérie Viet Triem Tong gave a talk on *Advanced threat representation*, at the CyberSchool Research School in 2024.

Valérie Viet Triem Tong participated to a round-table discussion on the topic of *cyber security training*, at the European Cyber Week (ECW) in 2024.

Pierre-François Gimenez participated to a round-table discussion on the topic of "Future vision for cybersecurity in Brittany and Europe: what challenges and what place for AI?" at the European Cyber Week (ECW) in 2024.

Pierre-François Gimenez participated to a round-table discussion on the topic of AI for defense at the "Séminaire IA du Commandement de la cyberdéfense".

Pierre-François Gimenez gave an invited talk ("Can generative AI help us better assess security solutions?") for the Journées Informatiques en Région Centre 2024, at INSA Centre Val de Loire.

Pierre-François Gimenez gave a talk on "Leveraging explainability to increase the usability of intrusion detection systems" at the Séminaire Sci-Rennes at Centre Inria de l'Université de Rennes.

On the Youtube page of the PIRAT team, many scientific talks are published. Most of them are recordings from the biweekly PIRAT seminars organized by Pierre-François Gimenez. In 2024, the channel reached 191 subscribers, with 61 published videos, about 13,000 views and 1000 hour of cumulated watch time.

Yufei Han gave a keynote talk on *Uncertainty in Cyber Security: Origin and Opportunities*, at the 3rd Workshop on Uncertainty Reasoning and Quantification in Decision Making, held in conjunction with ACM SIGKDD 2024.

Yufei Han gave an invited tutorial talk on *Backdoor Threats against Federated Learning Models: History and Open Problems*, at the 6th Workshop on Machine Learning for Cyber Security (MLCS), held in conjection with ECML PKDD 2024.

10.3.1 Specific official responsibilities in science outreach structures

Ludovic Mé is in charge of the cybersecurity program of the "Algorithms, Software and Usages" program Agency, which is supported by Inria for the entire French academic community.

Ludovic Mé organized and led the foresight seminar of the Inria theme"Security and Privacy" (Inria's teams working on the cybersecurity field, except cryptography).

Ludovic Mé serves:

- the steering committee off the PEPR cybersecurity (Programmes et Equipements Prioritaires de Recherche);
- the technical committee of the PTCC (Programme de Transfert au Campus Cyber);
- the Scientific Council of the LIRIMA (Laboratoire International de Recherche en Informatique et Mathématiques Appliquées);
- the Expert Council of the DSTN (Digital Science and Technology Network).

10.3.2 Participation in Live events

The PIRAT team has participed to the Breizh CTF 2024 by elaborating a sponsor challenge *CasinoLimit*, funded by Inria. The challenge has been procoposed to 120 teams of players for 600 persons. It has been deployed during 24h in OVH cloud and entirely solved by 5 teams. The challenge is available on Pirat public gitlab

Valérie Viet Triem Tong gave a talk on the theme of *Advanced Persistant Threat* at the conference *Who Run the Tech* organized by ESTIMnumérique.

10.3.3 Others science outreach relevant activities

Jean-Francois Lalande has participated to

- the program *1 scientifique, 1 classe : Chiche !* for the high school of Saumur.
- the juries of *15e Olympiades des Sciences de l'Ingénieur*.

Valérie Viet Triem Tong has participated to the program *1 scientifique, 1 classe : Chiche !* for the high school of Liffré.

Hélène Orsini has participated to the program *L codent L créent* in Collège des Gayeulles.

Gilles Guette has participated to the program *À la découverte de la recherche*. Gilles Guette has participated to the program "Plan National de Formation, Rendez vous de l'informatique : atelier réseaux" a training workshop for high school SNT/SNI teachers

11 Scientific production

11.1 Major publications

[1] X. Lyu, Y. Han, W. Wang, J. Liu, Y. Zhu, G. Xu, J. Liu and X. Zhang. 'Lurking in the shadows: Unveiling Stealthy Backdoor Attacks against Personalized Federated Learning'. In: Usenix Security 2024 - 33rd USENIX Security Symposium. Philadelphia,, United States, 2024, pp. 1–19. URL: https://inria.hal.science/hal-04827820.

11.2 Publications of the year

International journals

[2] X. Lyu, Y. Han, W. Wang, J. Liu, B. Wang, K. Chen, Y. Li, J. Liu and X. Zhang. 'CoBA: Collusive Backdoor Attacks with Optimized Trigger to Federated Learning'. In: *IEEE Transactions on Dependable and Secure Computing* (19th Aug. 2024), pp. 1–12. DOI: 10.1109/TDSC.2024.3445637. URL: https://inria.hal.science/hal-04829828 (cit. on p. 8).

[3] X. Xu, W. Wang, Z. Chen, B. Wang, C. Li, L. Duan, Z. Han and Y. Han. 'Finding the PISTE: Towards Understanding Privacy Leaks in Vertical Federated Learning Systems'. In: *IEEE Transactions on Dependable and Secure Computing* (2024), pp. 1–14. DOI: 10.1109/TDSC.2024.3445600. URL: https://inria.hal.science/hal-04829817 (cit. on p. 7).

International peer-reviewed conferences

[4] E. Anceaume, D. Frey and A. Rauch. 'Brief: Sharding in permissionless systems in presence of an adaptive adversary'. In: *31st International Colloquium on Structural Information and Communication Complexity (SIROCCO)*. 31st International Colloquium on Structural Information and Communication Complexity (SIROCCO). Vol. 14662. Lecture Notes in Computer Science. Vietri sul Mare, Italy, 23rd May 2024, pp. 481–487. DOI: 10.1007/978-3-031-60603-8_26. URL: https://hal.science/hal-04477243 (cit. on p. 9).

[5] E. Anceaume, D. Frey and A. Rauch. 'Sharding in permissionless systems in presence of an adaptive adversary'. In: *Proceedings of the International Conference on Networked Systems (NETYS)*. NETYS 2024 - 12th International Conference on Networked Systems. Rabat, Morocco: Springer, 2024, pp. 1–30. URL: https://cnrs.hal.science/hal-04794826 (cit. on p. 9).

[6] L. Cornanguer and P.-F. Gimenez. 'TADAM: Learning Timed Automata from Noisy Observations'. In: *Proceedings of the 2025 SIAM International Conference on Data Mining (SDM)*. SIAM International Conference on Data Mining (SDM25). Alexandria Virginia, United States, 1st May 2025. URL: https://hal.science/hal-04886774 (cit. on p. 8).

[7] P.-F. Gimenez and J. Mengin. 'Learning Conditional Preference Networks: an Approach Based on the Minimum Description Length Principle'. In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, 3rd-9th August 2024, Jeju, South Korea*. IJCAI 2024 - 33rd International Joint Conference on Artificial Intelligence. Jeju, South Korea: International Joint Conferences on Artificial Intelligence Organization, 2024, pp. 3395–3403. DOI: 10.24963/ijcai.2024/376. URL: https://ut3-toulouseinp.hal.science/hal-04572196.

[8] X. Lyu, Y. Han, W. Wang, J. Liu, Y. Zhu, G. Xu, J. Liu and X. Zhang. 'Lurking in the shadows: Unveiling Stealthy Backdoor Attacks against Personalized Federated Learning'. In: Usenix Security 2024 - 33rd USENIX Security Symposium. Philadelphia,, United States, 2024, pp. 1–19. URL: https://inria.hal.science/hal-04827820 (cit. on p. 7).

[9] X. Lyu, Y. Han, W. Wang, H. Qian, I. Tsang and X. Zhang. 'Cross-Context Backdoor Attacks against Graph Prompt Learning'. In: KDD 2024 : 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona, Spain: ACM, 2024, pp. 2094–2105. DOI: 10.1145/3637528.3671956. URL: https://inria.hal.science/hal-04827883.

[10] J.-M. Mineau and J.-F. Lalande. 'Evaluating the Reusability of Android Static Analysis Tools'. In: ICSR 2024 - 21st International Conference on Software and Systems Reuse. Vol. LNCS 14614. LNCS. Limassol, Cyprus: Springer, 2024, pp. 153–170. DOI: 10.1007/978-3-031-66459-5_10. URL: https://centralesupelec.hal.science/hal-04557993 (cit. on p. 6).

[11] H. Orsini and Y. Han. 'DYNAMO: Towards Network Attack Campaign Attribution via Density-Aware Active Learning'. In: SECRYPT 2024 - 21st International Conference on Security and Cryptography. Dijon, France: SCITEPRESS - Science and Technology Publications, 2024, pp. 91–102. DOI: 10.5220/0012759100003767. URL: https://inria.hal.science/hal-04877620 (cit. on p. 7).

[12] M. Poisson, R. Carnier and K. Fukuda. 'GothX: a generator of customizable, legitimate and malicious IoT network traffic'. In: *Cyber Security Experimentation and Test Workshop, CSET*. CSET - 17th Cyber Security Experimentation and Test Workshop. Vol. 17. Philadelphia, United States, 13th Aug. 2024, pp. 1–9. DOI: 10.1145/3675741.3675753. URL: https://inria.hal.science/hal-04629350 (cit. on pp. 6, 11).

[13] A. Schoen, G. Blanc, P.-F. Gimenez, Y. Han, F. Majorczyk and L. Me. 'A tale of two methods: unveiling the limitations of GAN and the rise of bayesian networks for synthetic network traffic generation'. In: *EuroS&PW 2024 : IEEE European Symposium on Security and Privacy Workshops*. 2024 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). Vienna, Austria: IEEE, 20th Aug. 2024, pp. 273–286. DOI: 10.1109/EuroSPW61312.2024.00036. URL: https://hal.science/hal-04871298 (cit. on p. 11).

[14] N. Talon, V. Viet Triem Tong, G. Guette, Y. Han and Y. Laarouchi. 'SCWAD: Automated Pentesting of Web Applications'. In: SECRYPT 2024 - 21st International Conference on Security and Cryptography. Dijon, France: SCITEPRESS - Science and Technology Publications, 2024, pp. 424–433. DOI: 10.5220/0012721000003767. URL: https://inria.hal.science/hal-04874868 (cit. on p. 9).

[15] Y. Zhou, Y. Han, H. Zhuang, H. Bao and X. Zhang. 'Attack-free Evaluating and Enhancing Adversarial Robustness on Categorical Data'. In: ICML 2024 - Forty-First International Conference on Machine Learning. Vienna, Austria, 2024, pp. 1–30. URL: https://inria.hal.science/hal-04827848 (cit. on p. 9).

National peer-reviewed Conferences

[16] F. Dijoud, P.-F. Gimenez, M. Hurfin, F. Majorczyk and B. Pilastre. 'Survey on system-level graph-based and anomaly-based intrusion detection'. In: RESSI 2024 - Rendez-Vous de la Recherche et de l'Enseignement de la Sécurité des Systèmes d'Information. Eppe-Sauvage, France, 2024, pp. 1–2. URL: https://inria.hal.science/hal-04714325.

Conferences without proceedings

[17] G. Quétel, E. Alata, P.-F. Gimenez, L. Pautet and T. Robert. 'A Parser-Based Data Collector for Intrusion Detection'. In: RESSI 2024 - Rendez-Vous de la Recherche et de l'Enseignement de la Sécurité des Systèmes d'Information. Eppe-Sauvage, France, May 2024, pp. 1–2. URL: https://hal.science/hal-04871463.

Reports & preprints

[18] T. Albouy, E. Anceaume, D. Frey, M. Gestin, A. Rauch, M. Raynal and F. Taïani. *Asynchronous BFT Asset Transfer: Quasi-Anonymous, Light, and Consensus-Free*. 15th May 2024. URL: https://inria.hal.science/hal-04578985 (cit. on p. 10).

[19] L. Miller, D. Pacaud, N. Derousseaux, E. Anceaume and R. Ludinard. *Mining in Logarithmic Space with Variable Difficulty*. 1st Nov. 2024. URL: https://cnrs.hal.science/hal-04794763 (cit. on p. 10).

Other scientific publications

[20] M. Poisson, S. Kilian, V. Viet Triem Tong, J.-F. Lalande, G. Guette, F. Ghihéry and D. Crémilleux. 'Digital twin for security evaluation using an attack agent: Discover attack paths in an information system without affecting the services in production.' In: USENIX 2024 - 33rd USENIX Security Symposium. Philadelphia, United States, 15th Aug. 2024, pp. 1–1. URL: https://inria.hal.science/hal-04708718.