Current Research at a Glance

Poster Collection

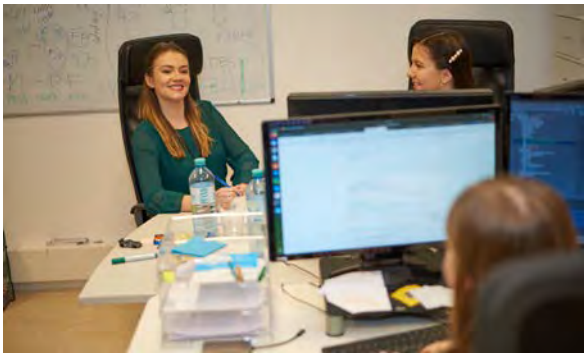


Table of Contents

10 Selected Top Papers

AREA 1 – Networked Systems Security

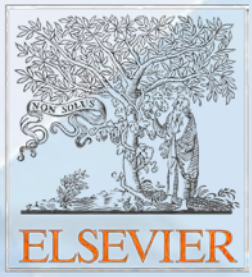
AREA 2 – Software Security

AREA 3 – Privacy and Secure Societies

AREA 4 – Applied Discrete Mathematics for Information Security

Beyond Research

10 Selected Top Papers



2017 IEEE European Symposium on Security and Privacy

Block Me If You Can: A Large-Scale Study of Tracker-Blocking Tools

Georg Merzdovnik*, Markus Huber†, Damjan Buhov*, Nick Nikiforakis‡,
Sebastian Neuner*, Martin Schmiedecker*, Edgar Weippl*
*SBA Research, {gmerzdovnik, dbuhov, sneuner, mschiedecker, eweippl}@sba-research.org
†St. Pölten UAS, markus.huber@fhstp.ac.at
‡Stony Brook University, nick@cs.stonybrook.edu

***Abstract*—In this paper, we quantify the effectiveness of third-party tracker blockers on a large scale. First, we analyze the architecture of various state-of-the-art blocking solutions and discuss the advantages and disadvantages of each method. Second, we perform a two-part measurement study on the effectiveness of popular tracker-blocking tools. Our analysis quantifies the protection offered against trackers present on more than 100,000 popular websites and 10,000 popular Android applications. We provide novel insights into the ongoing arms race between trackers and developers of blocking tools as well as which tools achieve the best results under what circumstances. Among others, we discover that rule-based browser extensions outperform learning-based ones, trackers with smaller footprints are more successful at avoiding being blocked, and CDNs pose a major threat towards the future of tracker-blocking tools.**

Overall, the contributions of this paper advance the field of web privacy by providing not only the largest study to date on the effectiveness of tracker-blocking tools, but also by highlighting the most pressing challenges and privacy issues of third-party tracking.

1. Introduction

In the modern internet, it has become a common practice for websites and mobile applications to rely on services provided by third parties. These services include advertisements, analytics, social integration widgets, and CDN-residing versions of popular JavaScript libraries. While the benefits of this third-party content integration are clear for the developers of first-party sites, the widespread adoption of these services is also inevitably linked with increased user tracking.

Every time a user's browser is instructed to fetch a third-party resource, that third-party server is given the ability to deliver tracking scripts and associate the first-party website with the bearer of third-party cookies and browser fingerprints. This tracking of online behavior allows for the construction of increasingly detailed user profiles, including sensitive information such as a user's political views and medical history. In addition to the exposure of users' online behavior to third parties, this third-party communication, which is typically unencrypted, can be further exploited by rogue ISPs and state-level attackers. For instance, it became publicly known that the National Security Agency (NSA) is piggybacking on third-party tracking cookies to de-anonymize Tor users and to identify targets for further exploitation [1], [2].

Third-party tracking thus has serious implications for the overall privacy and security of Internet users. Previous research focused on measuring the prevalence of tracking on common websites [3], [4], [5] and showed how privacy-conscious users and online trackers are at an arms race, with the former deleting their cookies and utilizing client-side, privacy-enhancing technologies, whereas the latter are migrating from traditional stateful tracking to more opaque, stateless tracking technologies based on browser fingerprinting [6], [7], [8], [9].

The absence of explicit policies regarding what a website is and is not allowed to do — coupled with the difficulty of setting and preserving opt-out cookies [10], [11], and the fact that the Do-Not-Track HTTP header is typically ignored by websites [3], [12], [13] — has motivated most savvy users to rely on client-side tools to preserve their online privacy. These client-side tools typically come in the form of browser extensions which differentiate between tracking and non-tracking HTTP requests, blocking the former and allowing the latter. At the time of writing, the two most common blocking tools are *AdBlock Plus* and *Ghostery*. AdBlock Plus focuses on blocking online advertisements, while Ghostery provides feedback on trackers included in websites. Note that even though advertisers do not necessarily need to track user interests in order to show ads, the majority of modern advertisers utilize Online Behavior Advertising which relies on building detailed profiles of a user's interests and is thus one more form of tracking. It is also worthwhile to note that some browser vendors such as Mozilla and Apple have recently acknowledged the importance of tracker-blocking tools and provide native support for rule-based blocking in their browsers [14], [15].

Despite the prevalence of these tracker-blocking tools, there is currently a lack of understanding of their effectiveness and applicability in the wild, and the extent to which they can protect users against motivated trackers. Previous research on the effectiveness of tracker-blocking tools is limited, both in scope as well as their considered threat models [13], [14], [16], [17]. In order to help close that gap,

DOI 10.1109/EuroSP.2017.26

Echoes of the Past: Recovering Blockchain Metrics From Merged Mining

Nicholas Stifter[1,3] ✉, Philipp Schindler[1], Aljosha Judmayer[1], Alexei Zamyatin[2,1], Andreas Kern[1], Edgar Weippl[1,3]

[1]SBA Research, Vienna, Austria
[2]Imperial College London, United Kingdom
[3]Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle (CDL-SQI), Institute of Information Systems Engineering, TU Wien
Email: (firstletterfirstname)(lastname)@sba-research.org

Abstract. So far, the topic of *merged mining* has mainly been considered in a security context, covering issues such as mining power centralization or cross-chain attack scenarios. In this work we show that key information for determining blockchain metrics such as the *fork rate* can be recovered through data extracted from merge mined cryptocurrencies. Specifically, we reconstruct a long-ranging view of forks and stale blocks in Bitcoin from its merge mined child chains, and compare our results to previous findings that were derived from live measurements. Thereby, we show that live monitoring alone is not sufficient to capture a large majority of these events, as we are able to identify a non-negligible portion of stale blocks that were previously unaccounted for. Their authenticity is ensured by cryptographic evidence regarding both, their position in the respective blockchain, as well as the Proof-of-Work difficulty.
Furthermore, by applying this new technique to Litecoin and its child cryptocurrencies, we are able to provide the first extensive view and lower bound on the stale block and fork rate in the Litecoin network. Finally, we outline that a recovery of other important metrics and blockchain characteristics through merged mining may also be possible.

1 Introduction

In blockchain-based cryptocurrencies the *fork rate* is considered to be an essential metric to better gauge the performance, capacity, and health of the respective communication network [1], and may also help in estimating other aspects such as their security [2] or degree of decentralization [3]. Furthermore, the fork rate can be indicative of adversarial behavior, such as *selfish mining* and its variants [2, 4–6] and other attacks that induce a higher ratio of stale blocks [7–9], or highlight periods of contention over protocol rule changes [10]. Historic and long-ranging data on stale blocks and the fork rate could also help determine the effectiveness of improvement measures and also provide a vital empirical basis for both predicting and directing future development.

However, for many cryptocurrencies such extensive data sets are not always readily available as a consequence of both design decisions, as well as the necessity to perform ongoing live monitoring to try and capture these events from gossip in the peer-to-peer (p2p) network. Moreover, while public sources of live monitoring data from popular

Published in ACM Computing Surveys 2019
Full Paper

From hack to elaborate technique - A survey on binary rewriting

MATTHIAS WENZL, FH Technikum Wien, Austria
GEORG MERZDOVNIK, SBA Research, Austria
JOHANNA ULLRICH and EDGAR WEIPPL, SBA Research, Austria and CDL-SQI, TU Wien, Austria

Binary rewriting is changing the semantics of a program without having the source code at hand. It is used for diverse purposes such as emulation (e.g., QEMU), optimization (e.g., DynInst), observation (e.g., Valgrind) and hardening (e.g., Control flow integrity enforcement). This survey gives detailed insight into the development and state-of-the-art in binary rewriting by reviewing 67 publications from 1966 up to 2018. Starting from these publications we provide an in-depth investigation of the challenges and respective solutions to accomplish binary rewriting. Based on our findings we establish a thorough categorization of binary rewriting approaches with respect to their use-case, applied analysis technique, code-transformation method and code generation techniques. We contribute a comprehensive mapping between binary rewriting tools, applied techniques and their domain of application. Our findings emphasize that although much work has been done over the last decades, most of the effort was put into improvements aiming at rewriting general purpose applications, but ignoring other challenges like altering throughput-oriented programs, or software with real-time requirements, that are often used in the emerging field of the Internet of Things. To the best of our knowledge, our survey is the first comprehensive overview on the complete binary rewriting process.

CCS Concepts: • **Software and its engineering → Software post-development issues**; *Automated static analysis*; *Dynamic analysis*; • **Security and privacy** → *Software and application security*.

Additional Key Words and Phrases: Binary rewriting, Binary hardening, Static rewriting, Dynamic rewriting , Minimal-invasive, Full-translation, Reassembly

ACM Reference Format:
Matthias Wenzl, Georg Merzdovnik, Johanna Ullrich, and Edgar Weippl. 2019. From hack to elaborate technique - A survey on binary rewriting. *ACM Comput. Surv.* 52, 3, Article 49 (June 2019), 36 pages. https://doi.org/10.1145/3316415

1 OVERVIEW

"Binary rewriting" describes the alteration of a compiled and possibly (dynamically) linked program without having the source code at hand in such a way that the binary under investigation stays executable [81]. Originally, binary rewriting was motivated by the need to change parts of a program during execution (e.g., run-time patching on the PDP-1 in the 1960's) [92]. Today, binary rewriting has evolved from a hack [92] through a repeatable technique for special purposes like link-time code optimization [5, 62] and performance optimization of win32 programs [112] to a plethora of approaches with applications in multiple domains. Popular applications are:

We express our gratitude to our reviewers who greatly helped to improve the paper with their valuable remarks and excavation of some early work references regarding binary rewriting.
Authors' addresses: Matthias Wenzl, FH Technikum Wien, Hoechstaedplatz 6, Vienna, 1200, Austria, wenzl@technikum-wien.at; Georg Merzdovnik, SBA Research, Favoritenstrasze 16, Vienna, 1040, Austria, gmerzdovnik@sba-research.org; Johanna Ullrich; Edgar Weippl, SBA Research, Favoritenstrasze 16, Vienna, 1040, Austria, CDL-SQI, TU Wien, Austria, JUllrich@sba-research.org, eweippl@sba-research.org.

Preprint: to appear at ACSAC 2017, December 4-8 2017

Grid Shock: Coordinated Load-Changing Attacks on Power Grids

The Non-Smart Power Grid is Vulnerable to Cyber Attacks as Well

Adrian Dabrowski
SBA Research
Wien, Austria
adabrowski@sba-research.org

Johanna Ullrich
SBA Research
Wien, Austria
jullrich@sba-research.org

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

ABSTRACT

Electric power grids are among the largest human-made control structures and are considered as critical infrastructure due to their importance for daily life. When operating a power grid, providers have to continuously maintain a balance between supply (i.e., production in power plants) and demand (i.e., power consumption) to keep the power grid's nominal frequency of 50 Hz or alternatively 60 Hz. Power consumption is forecast by elaborated models including multiple parameters like weather, season, and time of the day; they are based on the premise of many small consumers averaging out their energy consumption spikes.

In this paper, we develop attacks violating this assumption, investigate their impact on power grid operation, and assess their feasibility for today's adversaries. In our scenario, an adversary builds (or rents) a botnet of zombie computers and modulates their power consumption, e.g., by utilizing CPU, GPU, hard disks, screen brightness, and laser printers in a coordinated way over the Internet. Outperforming the grid's countervailing mechanisms in time, the grid is pushed into unstable states triggering automated load shedding or tie-line tripping. We show that an adversary does not have to rely on smart grid features to modulate power consumption given that an adequate communication infrastructure for striking the (legacy) power grid is currently nearly omnipresent: the Internet to whom more and more power-consuming devices are connected.

Our simulations estimate that between 2.5 and 9.8 million infections are sufficient to attack the European synchronous grid – depending on the mix of infected devices, the current mix of active power plant types, and the current overall produced power. However, the herein described attack mechanisms are not limited to the European grid.

ACM Reference Format:
Adrian Dabrowski, Johanna Ullrich, and Edgar R. Weippl. 2017. Grid Shock: Coordinated Load-Changing Attacks on Power Grids: The Non-Smart Power Grid is Vulnerable to Cyber Attacks as Well. In *2017 Annual Computer Security Applications Conference.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3134600.3134639

ACSAC 2017, December 4–8, 2017, San Juan, PR, USA

ACM ISBN 978-1-4503-5345-8/17/12…$15.00
https://doi.org/10.1145/3134600.3134639

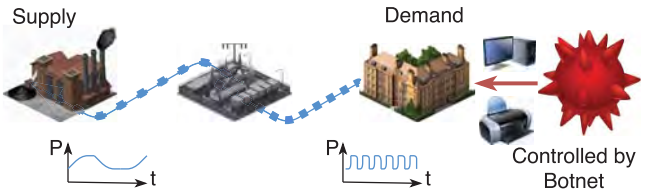


Figure 1: Visualization of Attacks 1 and 2: The botnet can modulate the power demand much faster than power plants can react.

1 INTRODUCTION

Electric power grids are among the largest human-made structures and by far the most important for technology-dependent societies. Without electricity, life as we know it would not function; there would be breakdowns in water and food supply, transport, medical aid, and communication infrastructures. For this reason, power grids are considered critical infrastructures, and operated with a high level of care to provide qualitative service, i.e., constant voltage and frequency. At the same time, power grids are legacy systems pre-dating modern telecommunication networks — such as the Internet — by decades, as is reflected in its structure: Electricity consumers are predominantly uncontrolled, i.e., consuming electric power whenever they need thereby causing fluctuations in consumption. However, on a macro scale fluctuations average out: for each consumer turning a light bulb off there is most likely another one turning the light on. Energy suppliers have developed sophisticated models that reliably forecast power demand in dependence of time of the day, week day, season and many other parameters allowing (centralized) power plants to trace actual consumption best possible in order to keep the equilibrium of production and consumption; the remaining gap is placed at disposal by so called *control reserves* (*spinning reserve* in the U.S.), i.e., the activation of power plants in stand-by.

Power grids around the globe currently undergo substantial modifications commonly summarized under the term *smart grid*, and the included concepts put an end to the strict separation of controlled production and uncontrolled consumption. On the one hand, renewables like wind turbines and photovoltaics provide electric energy in dependence of weather conditions and are thus only to a certain extent predictable, not to mention arbitrarily controllable. On the other hand, demand-side management aims to shift certain types of consumption, e.g., heating or cooling, in time. Synchronized over a communication channel, energy should then be consumed at the time of production by renewables. Due to such remote control of high amounts of power consumption, the smart grid is considered to be vulnerable to direct cyber attacks aiming to destabilize the system [28, 64].

Published at 26th USENIX Security Symposium 2017
Full Paper

"I Have No Idea What I'm Doing" - On the Usability of Deploying HTTPS

Katharina Krombholz, Wilfried Mayer, Martin Schmiedecker, and Edgar Weippl, *SBA Research*

https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/krombholz

This paper is included in the Proceedings of the 26th USENIX Security Symposium

August 16–18, 2017 • Vancouver, BC, Canada

ISBN 978-1-931971-40-9

Open access to the Proceedings of the 26th USENIX Security Symposium is sponsored by USENIX

Published at IEEE Symposium on Security and Privacy 2019
Full Paper

"If HTTPS Were Secure, I Wouldn't Need 2FA"- End User and Administrator Mental Models of HTTPS

Katharina Krombholz
CISPA Helmholtz Center for Information Security

Karoline Busse
Bonn University

Katharina Pfeffer
SBA Research

Matthew Smith
Bonn University
FhG FKIE

Emanuel von Zezschwitz
Bonn University
FhG FKIE

***Abstract*—HTTPS is one of the most important protocols used to secure communication and is, fortunately, becoming more pervasive. However, especially the long tail of websites is still not sufficiently secured. HTTPS involves different types of users, e.g., end users who are forced to make security decisions when faced with warnings or administrators who are required to deal with cryptographic fundamentals and complex decisions concerning compatibility.**

In this work, we present the first qualitative study of both end user and administrator mental models of HTTPS. We interviewed 18 end users and 12 administrators; our findings reveal misconceptions about security benefits and threat models from both groups. We identify protocol components that interfere with secure configurations and usage behavior and reveal differences between administrator and end user mental models.

Our results suggest that end user mental models are more conceptual while administrator models are more protocol-based. We also found that end users often confuse encryption with authentication, significantly underestimate the security benefits of HTTPS. They also ignore and distrust security indicators while administrators often do not understand the interplay of functional protocol components. Based on the different mental models, we discuss implications and provide actionable recommendations for future designs of user interfaces and protocols.

I. Introduction

In the context of information technologies, protecting communication content at large scale has become more important than ever before. Almost twenty years after Whitten and Tygar's usability evaluation of PGP [1], reliable encryption still cannot be taken for granted even though adoption rates are growing [2]. In today's Internet ecosystem, HTTPS is the fundamental cryptographic protocol to secure information in transit and to ensure data integrity and privacy between two communicating parties. However, HTTPS is still not the default for all websites, especially when it comes to the long tail of websites [2], [3]. At the time of writing, Internet-wide scans from SSLPulse suggest that 36,3% of sites surveyed still have inadequate security[1]. Recent studies, e.g., by Krombholz et al. [4], show that this is, among other reasons, due to the fact that the deployment of cryptographic protocols is a difficult task even for knowledgeable users. Similar to message encryption, HTTPS confronts different types of (mostly technically adept) users with cryptographic algorithms and protocols which they do not fully understand – see, e.g., Krombholz et al. [4], Green and Smith [5], Acer et al.[3], Fahl et al. [6], Oltrogge et al. [7], and Reeder et al. [8]. In addition, users who are exposed to poorly configured sites are forced to make security-critical decisions and are often not aware of the respective consequences.

We argue that we still do not understand *why* these carefully designed protocols do not meet the needs of (knowledgeable) users to securely operate cryptographic applications. Therefore, this work employs an inductive approach to learn about the root causes for user misconceptions by formalizing mental models of end users and administrators. In particular, we focus on how users think that HTTPS works and against which types of attackers they think they are protected. By doing so, we get a detailed understanding of which knowledge gaps have to be filled in future protocol designs. We thereby contribute a qualitative study with 18 end users and 12 experienced administrators; our findings reveal interesting differences in the mental models of these two distinct user groups.

We found that many non-expert participants significantly underestimate the level of protection that HTTPS offers, whereas administrators generally have a good understanding of what HTTPS can or cannot protect against. We also discovered that most administrators have little conceptual knowledge of how the protocol works but are very familiar with the different steps of establishing a communication. Key elements are often considered as blackboxes and poorly understood. We further found that the distinction between authentication and encryption is unclear to many users–even to some experts. Based on our findings, we identified protocol components that diverge from user mental models and discuss implications and potential countermeasures.

The goal of this paper is to derive and compare mental models in order to understand if and how they deviate from the underlying functionality of HTTPS and their impact on security. The main contributions of this paper are as follows:

We conducted an in-depth qualitative study with $n = 30$ participants to **formalize user mental models** and **threat models** and to **understand users' perceptions, attitudes and misconceptions of how HTTPS works**. By focusing on

[1]https://www.ssllabs.com/ssl-pulse/, Accessed: 10/30/2018

Investigating System Operators' Perspective on Security Misconfigurations

Constanze Dietrich
Berliner Hochschule für Technik*
constanze.die@gmail.com

Katharina Krombholz
CISPA Helmholtz Center (i.G.)†
krombholz@cispa.saarland

Kevin Borgolte
Princeton University‡
kevin@iseclab.org

Tobias Fiebig
TU Delft§
t.fiebig@tudelft.nl

ABSTRACT

Nowadays, security incidents have become a familiar "nuisance," and they regularly lead to the exposure of private and sensitive data. The root causes for such incidents are rarely complex attacks. Instead, they are enabled by simple misconfigurations, such as authentication not being required, or security updates not being installed. For example, the leak of over 140 million Americans' private data from Equifax's systems is among most severe misconfigurations in recent history: The underlying vulnerability was long known, and a security patch had been available for months, but was never applied. Ultimately, Equifax blamed an employee for forgetting to update the affected system, highlighting his *personal* responsibility.

In this paper, we investigate the operators' perspective on security misconfigurations to approach the human component of this class of security issues. We focus our analysis on system operators, who have not received significant attention by prior research. Hence, we investigate their perspective with an inductive approach and apply a multi-step empirical methodology: *(i)* a *qualitative* study to understand how to approach the target group and measure the misconfiguration phenomenon, and *(ii)* a *quantitative* survey rooted in the qualitative data. We then provide the first analysis of system operators' perspective on security misconfigurations, and we determine the factors that operators perceive as the root causes. Based on our findings, we provide practical recommendations on how to reduce security misconfigurations' frequency and impact.

CCS CONCEPTS

• **Security and privacy → Social aspects of security and privacy**; Usability in security and privacy; • **Social and professional topics → Employment issues**; **Computing occupations**;

KEYWORDS

Computer systems; system operations; operators; administrators; security; misconfigurations; vulnerabilities; human factors.

ACM Reference Format:
Constanze Dietrich, Katharina Krombholz, Kevin Borgolte, and Tobias Fiebig. 2018. Investigating System Operators' Perspective on Security Misconfigurations. In *2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18), October 15–19, 2018, Toronto, ON, Canada.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3243734.3243794

CCS '18, October 15–19, 2018, Toronto, ON, Canada

ACM ISBN 978-1-4503-5693-0/18/10…$15.00
https://doi.org/10.1145/3243734.3243794

1 INTRODUCTION

Security incidents and vulnerabilities in today's Internet are often believed to be caused by programming errors, such as faulty input validation, race conditions, or buffer overflows, that are exploited to disrupt services without the vulnerability being publicly known and before a patch is available (0 days). However, when investigating recent security incidents, such as those of Equifax [2, 3], we find a different picture. The vulnerability exploited in the primary Equifax incident, in which personally identifiable information of 143 million customers were inadvertently disclosed and which sparked a congressional inquiry, was clearly a programming mistake. However, while a patch to address the bug was released months prior, it was simply not yet deployed to the production environment.

Of course, not applying (security) patches can have its cause in countless reasons, such as technical debt accumulated over time, or availability and functionality requirements. Yet, when investigating the Equifax incident, such complex reasons are not the breach's cause. In the end, Equifax blamed the entire incident on a single operator for forgetting to install security patches in time [4].

Broadening the scope, incidents that have their root cause in human error can be found all over the Internet, from basic infrastructure to applications [5, 6]. For example, in early 2015, over 40,000 MongoDB instances were publicly accessible from the Internet, without authentication and authorization, and, in turn, allowed anyone to retrieve the stored data [7], which might have been confidential or possibly would have even required governmental security clearances. In fact, one of these MongoDB instances contained millions of voting records from Mexican citizens, and, in turn, it leaked them online [8]. Other database systems, like Redis or memcached, are not spared from similar human error: hundreds of thousands of systems were discovered to be unprotected [6]. The configuration of Transport Layer Security (TLS) for web application servers are often similarly vulnerable to misconfigurations due to human error [9]. Ultimately, misconfigurations can also lead to other vulnerabilities, such as servers becoming vulnerable to denial-of-service attacks [10, 11], or websites turning malicious [12] or being defaced to embarrass the systems' operators [13, 14].

The overarching aspect of these incidents is that the mistake leading to the incident occurred during the *operation* of the affected system instead of its *development* (as it is the case for software vulnerabilities). These mistakes do not need to be complex, but they can even be comparatively simple errors, such as missing or incorrect

*We use "Berliner Hochschule für Technik" instead of "Beuth Hochschule für Technik Berlin" because of Peter Christian Wilhelm Beuth's antisemitic views [1]. We stand for a diverse and inclusive scientific community, and we do not want to perpetuate the name of a researcher who did not.
†Research partially performed while at SBA Research.
‡Research performed while at UC Santa Barbara.
§Research partially performed while at TU Berlin.

Published in Theoretical Computer Science 2019
Full Paper

Theoretical Computer Science 800 (2019) 90–106

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Problems and algorithms for covering arrays via set covers

Ludwig Kampel, Manuel Leithner, Bernhard Garn, Dimitris E. Simos *

SBA Research, Vienna A-1040, Austria

ARTICLE INFO

Article history:
Available online 15 October 2019

Keywords:
Covering arrays
Set covers
Weighted budgeted covering arrays
Weighted budgeted set covers
Greedy heuristic algorithms

ABSTRACT

In this paper, we explore some connections between covering arrays (CAs) and set covers (SCs) that already existed in the literature, and in some cases we provide new mappings between these structures. In particular, the devised mappings make feasible an interpretation of weighted budgeted CAs (WBCAs) as weighted budgeted SCs. These connections in turn make it possible to reformulate known greedy heuristics for computing mixed-level CAs and evolve new algorithms for WBCA generation. This also enables importing an upper bound on the size or a lower bound on the covered weight of the generated arrays. We further carry out a comparison of a CA generation strategy that has an analogue in the SC world with one developed specifically for CAs.
Moreover, we experiment with several problem instances for CA and WBCA generation, and compare CA solvers versus SC solvers, both in quality of their output size and covered weights, as well as computation time. Our experiments underpin the hypothesis that CA solvers provide solutions of comparable quality to the ones returned by the considered SC solvers, although the latter solvers generally provide solutions of better quality. Nevertheless, CA solvers provide solutions to the respective problem instances much faster.

1. Introduction

Covering arrays (CAs) are discrete structures appearing in design theory. Most frequently, they are introduced as arrays having specific coverage properties regarding the appearance of tuples. In this regard, they can be considered a generalization of orthogonal arrays. These covering properties make CAs very interesting for both theoretical and practical use. In recent years, CAs have attracted significant research attention due to their applications in automated software testing [23,24,32].

See [25] for a survey on CA generation methods and [11] for a survey on combination testing strategies. Despite the extensive efforts of numerous researchers, finding an optimal CA (i.e. a CA with minimal number of rows) for a given configuration remains a challenging problem. Few results exist regarding the covering array number, the smallest number of rows for which a certain CA exists. More precisely, this number is completely determined only in the case of binary CAs of strength two [22] and for alphabet sizes that are a prime power, where the number of columns is restricted with respect

* Corresponding author.
E-mail addresses: lkampel@sba-research.org (L. Kampel), mleithner@sba-research.org (M. Leithner), bgarn@sba-research.org (B. Garn), dsimos@sba-research.org (D.E. Simos).

https://doi.org/10.1016/j.tcs.2019.10.018
0304-3975/

The Other Side of the Coin: User Experiences with Bitcoin Security and Privacy*

Katharina Krombholz, Aljosha Judmayer, Matthias Gusenbauer, and Edgar Weippl

SBA Research, Vienna, Austria
(firstletterfirstname)(lastname)@sba-research.org

Abstract. We present the first large-scale survey to investigate how users experience the Bitcoin ecosystem in terms of security, privacy and anonymity. We surveyed 990 Bitcoin users to determine Bitcoin management strategies and identified how users deploy security measures to protect their keys and bitcoins. We found that about 46% of our participants use web-hosted solutions to manage at least some of their bitcoins, and about half of them use exclusively such solutions. We also found that many users do not use all security capabilities of their selected Bitcoin management tool and have significant misconceptions on how to remain anonymous and protect their privacy in the Bitcoin network. Also, 22% of our participants have already lost money due to security breaches or self-induced errors. To get a deeper understanding, we conducted qualitative interviews to explain some of the observed phenomena.

1 Introduction

With a current market capitalization of more than 3.5 billion USD, Bitcoin is the most successful cryptographic currency at this time. Bitcoin is utilized for roughly 130.000 transactions per day [6] and has gained significant news coverage. With the success of Bitcoin, several other cryptographic currencies were developed either based on Bitcoin or from scratch.

Although the popularity of cryptographic currencies is increasing, they are not yet a mass phenomenon. One of the reasons is that Bitcoin forces its users to deal with public key cryptography. Furthermore, Bitcoin shifts the responsibilities for most security measures to the end user compared to centralized monetary systems. Even though there is a great variety of software available for managing bitcoins, user-experience is still not obviating the need to deal with the technical fundamentals and to perform backups to recover their virtual monetary assets in case of a loss. Hence, these systems are not resilient to human errors. Reports from online forums and mailing-lists show that many Bitcoin users already lost money due to poor usability of key management and security breaches such as malicious exchanges and wallets. This motivates our research on human interactions with the Bitcoin ecosystem.

Bitcoin users have a huge variety of tools available to manage their virtual assets. These tools are commonly referred to as *wallets*. A wallet was originally defined as a collection of private keys [8]. Hence, a piece of paper with a private key on it or even a

* This is a pre-conference version of the paper to appear at *Financial Cryptography and Data Security 2016*

Published at IEEE Symposium on Security and Privacy 2019
Full Paper

XCLAIM: Trustless, Interoperable, Cryptocurrency-Backed Assets

Alexei Zamyatin[†‡], Dominik Harz[†], Joshua Lind[†], Panayiotis Panayiotou[†], Arthur Gervais[†], William Knottenbelt[†]
[†] Imperial College London [‡] SBA Research

***Abstract*—Building trustless cross-blockchain trading protocols is challenging. Centralized exchanges thus remain the preferred route to execute transfers across blockchains. However, these services require trust and therefore undermine the very nature of the blockchains on which they operate. To overcome this, several decentralized exchanges have recently emerged which offer support for atomic cross-chain swaps (ACCS). ACCS enable the trustless exchange of cryptocurrencies across blockchains, and are the only known mechanism to do so. However, ACCS suffer significant limitations; they are slow, inefficient and costly, meaning that they are rarely used in practice.**

We present XCLAIM: the first generic framework for achieving trustless and efficient cross-chain exchanges using *cryptocurrency-backed assets* (CBAs). XCLAIM offers protocols for issuing, transferring, swapping and redeeming CBAs securely in a non-interactive manner on existing blockchains. We instantiate XCLAIM between Bitcoin and Ethereum and evaluate our implementation; it costs less than USD 0.50 to issue an arbitrary amount of Bitcoin-backed tokens on Ethereum. We show XCLAIM is not only faster, but also significantly cheaper than atomic cross-chain swaps. Finally, XCLAIM is compatible with the majority of existing blockchains without modification, and enables several novel cryptocurrency applications, such as cross-chain payment channels and efficient multi-party swaps.

***Index Terms*—blockchain, interoperability, CBA, Bitcoin, Ethereum**

I. INTRODUCTION

Blockchain-based cryptocurrencies enable secure and trustless transactions between parties. As a result, they have gained widespread adoption and popularity in recent years; there are currently over 2 000 different cryptocurrencies in operation, with a total market cap of USD 135bn [48]. However, despite a growing and thriving ecosystem, cryptocurrencies continue to operate in complete isolation from one another: blockchain protocols provide no means by which to communicate or exchange data with external systems. Hence, achieving interoperability between blockchains remains an open challenge.

Centralized exchanges thus remain the preferred route to execute fund transfers and exchanges across blockchains. However, these services require trust and therefore undermine the very nature of the cryptocurrencies on which they operate, making them vulnerable to attacks [32], [35], [89], [95]. To overcome this, *decentralized exchanges* [1], [2], [13], [16], [18] (*DEXs*) have recently emerged, removing the need to trust centralized intermediaries for blockchain transfers. However, the vast majority of DEXs only enable the exchange of *cryptocurrency-assets* within a single blockchain, i.e., they do not operate across blockchains (*cross-chain*). As such, it is only a handful of platforms [17], [30] that actually support cross-chain exchanges through the use of *atomic cross-chain swaps* (ACCS) [4], [33], [69], [105].

ACCS enable secure cross-chain exchanges, e.g. using *hashed timelock contracts* (HTLCs) [5], [47]. At present, they are the only mechanism to do this without necessitating trust. Unfortunately, they require several strong assumptions to maintain security, thus limiting their practicality: they are interactive, requiring all parties to be online and actively monitor all involved blockchains during execution; they require synchronizing clocks between blockchains and rely on pre-established secure out-of-band communication channels. In addition, they also incur long waiting periods between transfers and suffer the limitation that for every cross-chain swap, four transactions need to occur, two on each blockchain. This makes them expensive, slow and inefficient.

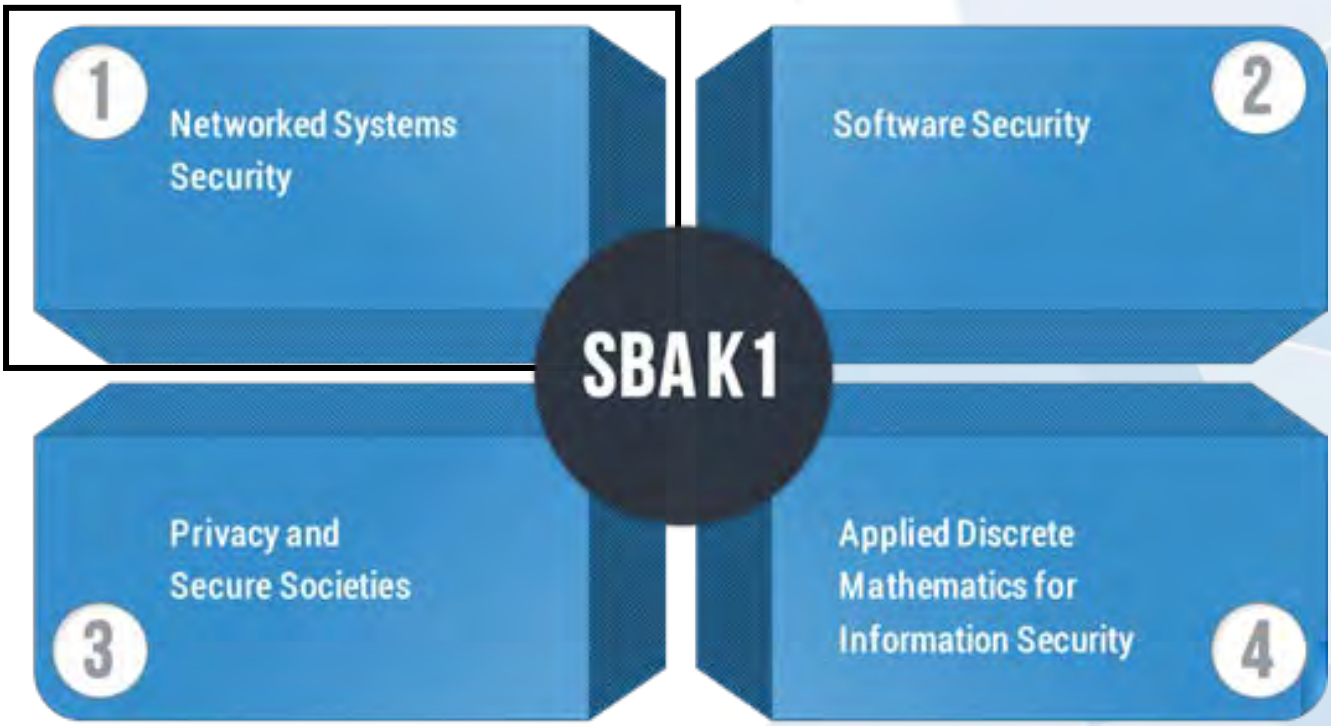
We therefore present XCLAIM (pronounced *cross-claim*): the first generic framework for achieving trustless cross-chain exchanges using *cryptocurrency-backed assets*. In XCLAIM, blockchain-based assets can be securely constructed and one-to-one backed by other cryptocurrencies, for example, Bitcoin-backed tokens on Ethereum. Through the secure issuance, swapping, and redemption of these assets, users can perform cross-chain exchanges in a trustless and non-interactive manner, overcoming the limitations of existing solutions.

To achieve this, XCLAIM exploits publicly verifiable *smart contracts* to remove the need for trusted intermediaries and leverages *chain relays* [6], [33], [76], [106] for cross-chain state verification. Using these building blocks, XCLAIM constructs a publicly verifiable audit log of user actions on both blockchains and employs *collateralization* and *punishments* to enforce the correct behavior of participants. Thereby, XCLAIM follows a *proof-or-punishment* approach, i.e., participants must proactively prove adherence to system rules.

Due to its simple and efficient design, XCLAIM enables several novel applications, such as: (i) *cross-chain payment channels*, where users can exchange payments *off-chain* across different blockchains in a trustless manner; (ii) *temporary transaction offloading*, where users temporarily tokenize their cryptocurrency on other blockchains to overcome network congestion and high fees; and (iii) *N-way and multi-party atomic swaps* allowing efficient and complex atomic swaps. Finally, as XCLAIM maintains compatibility with existing standardized asset interfaces [10], [11], the issued assets are tradeable via existing decentralized exchanges, enabling these

Networked Systems Security

AREA 1



Actively Probing Routes for Tor AS-level Adversaries with RIPE Atlas

Wilfried Mayer, Georg Merzdovnik, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

- Tor provides **anonymity** to millions of users.
- Low-latency anonymity systems are vulnerable to **traffic correlation attacks**.
- Strong passive adversaries, such as large **autonomous systems (AS)**.
- Current analyses mostly based on **BGP updates**.
- With **RIPE Atlas** and **traceroute** this needs to be re-evaluated.
- **Placement of measurement nodes** in the same ASes as Tor network nodes.

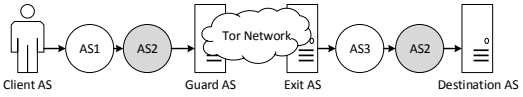


Figure 1: AS2 in a possible position for a **traffic correlation attack**

Tor Relays & RIPE Atlas

- Tor relays (approx. 6,500) are globally distributed.

	Relays	Diff. AS	BW (Gbit/s)
All Relays	6,509	1,104	418.07
Exit Relays	1,000	275	112.90
Guard Relays	2,415	470	254.61

Table 1: Tor relay overview

- RIPE Atlas measurement probes (approx. 10,000) are also globally distributed.
- Partially, in the same autonomous systems.

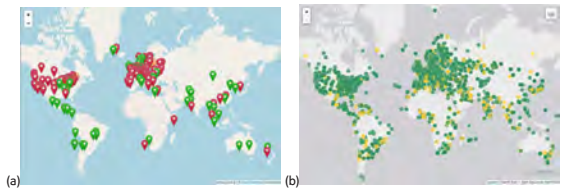


Figure 2: Visualization of Tor relays (a) and RIPE Atlas coverage (b)

- RIPE Atlas probes cover a substantial amount of ASes with Tor relays.
- For Tor exit relays it is 41% of total exit probability.
- For Tor guard relays it is 83% of total exit probability.

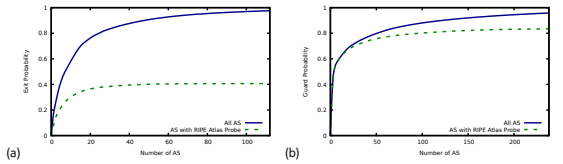


Figure 3: Accumulated percentage of (a) exit, and (b) guard probability with the number of autonomous systems

- Numbers could be increased:
  - Add 10 selected probes, cover 87% exit probability.

`traceroute` Measurements

- Executed `traceroute` commands on RIPE Atlas probes.
  - Placed in AS with Tor relays.
- Four different directions of measurements.
- From Top client ASes to Top destination ASes.

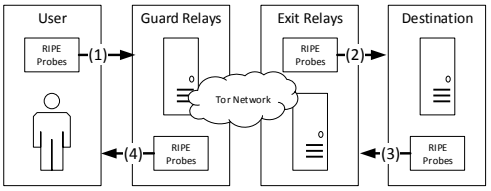


Figure 4: Four different directions of active RIPE Atlas traceroute scans

Evaluation

- Identified ASes with high probability to be on guard side as well as exit side (from single client AS to top destination ASes).

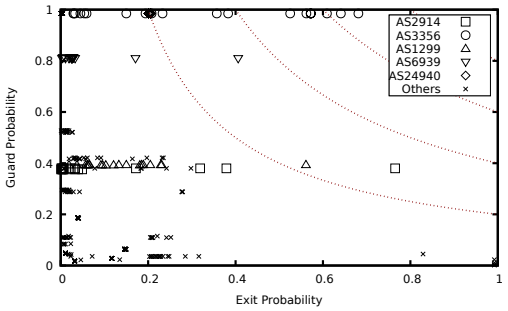


Figure 5: Combined probability of ASes appearing on the client and destination path

Conclusion

- A **novel way to analyze the network routes** taken by traffic from and to the Tor network.
- Utilized the **RIPE Atlas** framework.
- Identified a **small set of ASes which have a great influence** on the total amount of Tor bandwidth.
- We generated a **valuable additional data source**.

Wilfried Mayer, Georg Merzdovnik, and Edgar Weippl. Actively Probing Routes for Tor AS-level Adversaries with RIPE Atlas. In *IFIP International Information Security and Privacy Conference*, 2020.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Automated Emulation of IoT Device Firmware

Christian Kudera, Sebastian Dietz, Georg Merzdovnik

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Background & Motivation

Security in the IoT Ecosystem

Achieving security in the IoT ecosystem is a challenging task. According to Kaspersky [1], over 100 million attacks against the IoT were identified in the first half of 2019.

Compromised IoT devices are misused for:

- Distributed denial-of-service (DDoS) attacks
- Spamming and cryptocurrency mining
- Proxy agents or VPN pools

Problems

- The IoT market is rapidly growing, whereby the devices are characterized by heterogeneity due to different architectures and protocols [2].
- The emulation of IoT devices is rather limited, especially if the firmware doesn't contain a Linux operating system [3].
- Due to the lack of suitable tools, the security analysis of IoT devices is challenging, time consuming, and not well supported [4, 5].

Need: Framework that automatically builds emulated IoT devices from firmware samples without any further knowledge.

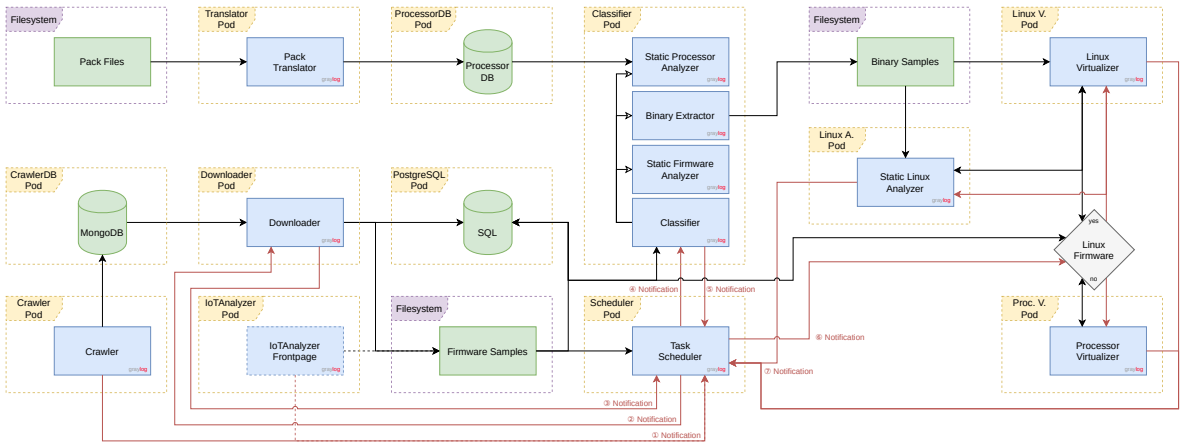


Figure 1: Automated Emulation of IoT Device Firmware Architecture

Methodology

Figure 1 illustrates the architecture of the framework and its corresponding components:

- **Crawler & Downloader**: The Internet is constantly monitored for new firmware samples. Identified samples are downloaded for an analysis.
- **Task Scheduler:** Coordinates the tasks for all components, with all intermediate results being stored in a database.
- **Classifier:** Classifies the samples and extracts as much information as possible through a modular approach. For example, the *Static Processor Analyzer* uses previously gained knowledge about different processors (e.g., base address, memory size) to determine the processor family and peripherals.
- **Processor & Linux Virtualizer:** Uses the knowledge gained from the classifier to create suitable virtualization instances. Lacking knowledge is complemented through dynamic analyses (e.g., identifying appropriate processors and peripherals).
- **Static Linux Analyzer:** Extracts additional knowledge from the Linux file system (e.g., installed software, password hashes).
- **IoT Analyzer Frontpage:** In the future, we will provide a website, where firmware samples can be uploaded. The firmware sample will be analyzed and the uploader receives a report.

Outlook

IoT Firmware Fuzzing & Symbolic Execution

- Dynamically probe the firmware for security issues
- Analyze the behaviour of the firmware in different scenarios

IoT Honeypots

- Collect IoT-relevant malicious empirical data
- Formulate IoT-centric attack signatures
- Generate IoT-specific technical threat intelligence

IoT Device Characteristics Database

- Manufacturer name and device model
- Open ports and running services
- Device fingerprints and application banners
- Device interaction information

Large-scale Identification of Exploited IoT Devices

- Use the gained knowledge to identify compromised IoT devices in the Internet
- Inform national and global CERT teams about ongoing threats

[1] IoT under fire: Kaspersky detects more than 100 million attacks on smart devices in H1 2019, 10 2019.
[2] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose. SoK: Security Evaluation of Home-Based IoT Deployments. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1362–1380, 2019.
[3] Z. K. Zhang, M. C. Y. Cho, C. W. Wang, C. W. Hsu, C. K. Chen, and S. Shieh. IoT Security: Ongoing Challenges and Research Opportunities. In *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, pages 230–234, Nov 2014.
[4] J. Wetzels. Ghost in the Machine: Challenges in Embedded Binary Security. Usenix Enigma, February 2017.
[5] M. Muench, J. Stijohann, F. Kargl, A. Francillon, and D. Balzarotti. What you corrupt is not what you crash: Challenges in fuzzing embedded devices. In *NDSS 2018, Network and Distributed Systems Security Symposium, 18-21 February 2018, San Diego, CA, USA*, 2018.

This research is funded by the FFG under the BRIDGE Young Scientists program (grant no. 872664) and the industrial PhD program (grant no. 878784).

Automated Security Risk Identification Based on Engineering Data

Matthias Eckhart, Andreas Ekelhart, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Background & Motivation

Security by Design for Industry 4.0

Security risk management efforts are vital for adopting a security-enhanced engineering process for cyber-physical systems [4], as security risks need to be addressed in a cost-effective manner.

This requires:

- Integration of security aspects in digitized engineering workflows
- Utilization of existing engineering know-how
- Adoption of the IEC 62443-3-2 [6] for security risk assessments
- Methods to efficiently identify risk sources and attack consequences

Problems

- Existing engineering data formats, such as AutomationML [2], lack semantic modeling concepts for expressing security know-how
- Carrying out security risk assessments according to IEC 62443-3-2 [6] is complex and effortful since tool support is missing
- Identifying and understanding relationships among the detected security risks is difficult [3]

Need: Method that automatically identifies security weaknesses based on engineering data and visualizes their cyber-physical relevance.

Automated Risk Identification Method

Contribution [5]: Provide a method for the automated identification of security risks based on AutomationML engineering artifacts and visualization of threats by means of cyber-physical attack graphs (CPAGs).

Overview

- Risk identification method follows the IEC 62443-3-2 [6]
- Introduction of a security modeling concept named *AMLsec* for the adequate representation of the cyber-physical systems' security properties in AutomationML [2] artifacts
- Knowledge-based approach
- Novel variant of attack graphs that systems integrators can apply to gain insights into possible multistage cyber-physical attacks
- Open-source prototype: `https://github.com/sbaresearch/amlsec`

Ontological Security Modeling

- Engineering knowledge present in AutomationML artifact is transformed to OWL
- Conceptual mapping via AutomationML libraries
- Data validation checks via reasoners and SHACL constraints
- ICS security ontology and SHACL shapes comprise security know-how
- Knowledge is interlinked with data from public sources (e.g., CVEs)

Automated Risk Identification

- Based on a combination of SPARQL queries and SHACL constraints
- SPARQL queries check whether zone and conduit requirements as per the IEC 62443-3-2 [6] are met
- Validation rules expose insecure components and configurations
- Interlinked knowledge graph (e.g., CVEs) reveals known vulnerabilities
- Attack consequences are identified by interpreting the semantics of plant components and associating them to assets that are at risk

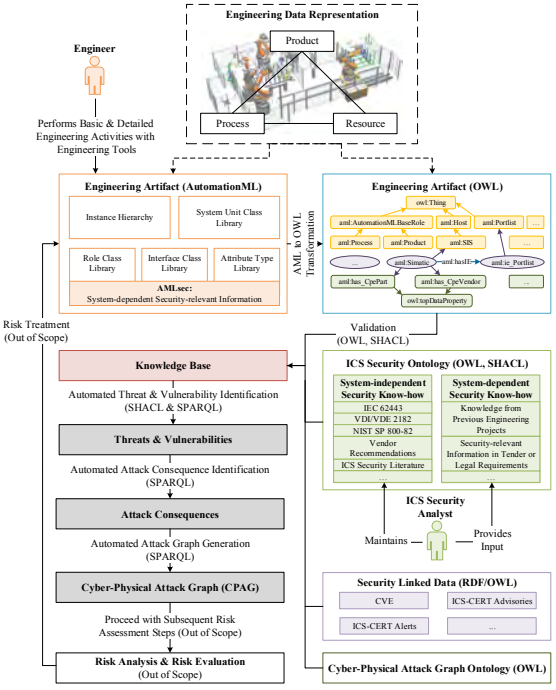


Figure 1: Overview of our AutomationML-based risk identification method (robot cell illustration taken from [1]).

Cyber-Physical Attack Graphs

A cyber-physical attack graph (CPAG) is a directed vertex- and edge-weighted graph $\boldsymbol{CPAG} = (\boldsymbol{V}, \boldsymbol{E}, \omega_{\boldsymbol{V}}, \omega_{\boldsymbol{E}})$, where $\boldsymbol{V}$ is the finite vertex set of assets, $\boldsymbol{E}$ is a multiset of directed edges from $\boldsymbol{V} \times \boldsymbol{V}$ representing vulnerabilities, $\omega_{\boldsymbol{V}}\colon \boldsymbol{V} \to \boldsymbol{S}$ is the vertex weight function that maps all vertices according to the assets' cyber-physical criticality onto the set $\boldsymbol{S}$, $\omega_{\boldsymbol{E}}\colon \boldsymbol{E} \multimap \boldsymbol{S}$ is the edge weight function that maps all edges according to the vulnerabilities' severity onto the set $\boldsymbol{S}$, and $\boldsymbol{S} = [0, 10]$.

CPAGs can be automatically generated by means of a SPARQL query. Subsequent pruning increases the utility and usability of CPAGs.

Conclusion & Outlook

- Method fosters a security-by-design engineering process
- Automated identification of risk sources and attack consequences
- Introduction of security concepts for AutomationML (AMLsec)
- Method seamlessly integrates into the engineering workflow
- CPAGs visualize potential multistage cyber-physical attacks that exploit the weaknesses identified in engineering data
- Open-source prototypical implementation of our method exists
- Incorporating COLLADA and PLCOpen XML is worthwhile
- Quantitative analysis capabilities will be added in the future

[1] AutomationML. AutomationML example: Robot cell. Technical report, March 2017.
[2] R. Drath, A. Luder, J. Peschke, and L. Hundt. AutomationML - the glue for seamless automation engineering. In *2008 IEEE International Conference on Emerging Technologies and Factory Automation*, pages 616–623, Sept 2008.
[3] M. Eckhart, B. Brenner, A. Ekelhart, and E. Weippl. Quantitative security risk assessment for industrial control systems: Research opportunities and challenges. *Journal of Internet Services and Information Security (JISIS)*, 9(3):52–73, August 2019.
[4] M. Eckhart, A. Ekelhart, A. Lüder, S. Biffl, and E. Weippl. Security development lifecycle for cyber-physical production systems. In *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, volume 1, pages 3004–3011, Oct 2019.
[5] M. Eckhart, A. Ekelhart, and E. Weippl. Automated security risk identification based on engineering data (in review). 2020.
[6] IEC. 62443-3-2: Security for industrial automation and control systems – part 3-2: Security risk assessment and system design. *International Standard, Draft, International Electrotechnical Commission, Geneva*, 1, 2018.

This research was further funded by the FFG under the industrial PhD program (grant no. 874644). Moreover, the financial support by the Christian Doppler Research Association, the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

AVRS: Emulating AVR Microcontrollers for Reverse Engineering and Security Testing

Michael Pucher, Christian Kudera, Georg Merzdovnik

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Background

AVR 8-bit microcontrollers are used in automotive applications, sensor nodes and IoT devices, while also being popular with hobbyists. They are the basis for most boards of the Arduino product line, which in turn is used in a variety of projects, like data logging [5] or control and measurement systems [3].

Due to weak processing power, security can become an afterthought:

- Data received from peripherals is untrusted user input.
- Simple vulnerabilities persist, e.g., stack buffer overflows.
- AVR devices cannot be monitored efficiently during execution.

Problems

- While there is a wealth of different tools for analyzing firmware for architectures that are common to standard hardware, such as Arm [2, 1], the AVR ecosystem did not receive a similar treatment.
- Muench et al. [2] have shown the need for full-system emulators for developing potent security analysis techniques.
- Existing emulators, e.g., Avrora [4], are intended to be used as part of the development process and thus unfit for use on unknown firmware.

Need: Emulators specifically designed for security analysis.

Methodology

- **Analysis of Existing Emulators.** Analyze existing open-source AVR emulators and document shortcomings
- **AVRS.** Implement a new emulator, compensating existing shortcomings
- **Evaluation of all Emulators.** Compare emulators regarding performance and completeness, examine AVRS improvements
- **Fuzzing AVR Firmware.** Build fuzzing support on top of AVRS

AVRS

Designed to reuse existing approaches of other emulators, improve on lacking monitoring features and performance

- **Instruction Decode.** Decode all instructions upfront, use internal intermediate representation (IR)
- **Instruction Emulation.** Emulating IR optimizes to jump table
- **Device Differences.** Architectural differences are decided at compile-time, using procedural Rust macros
- **Peripherals and Monitoring.** Provide Rust and RPC API for peripheral implementation, API can hook MMU for monitoring memory access.

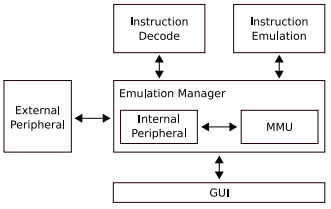


Figure 1: Overview of AVRS Architecture

Fuzzing

- **Implementation.** Use *boofuzz* as input generator, AVRS as emulator
- **Crash Detection.** Use heuristics presented in [2] to detect crashes
- **Example Programs.** Test fuzzer using serial protocol on ATtiny104 and ATmega328p, providing fuzzing cores and peripheral implementations

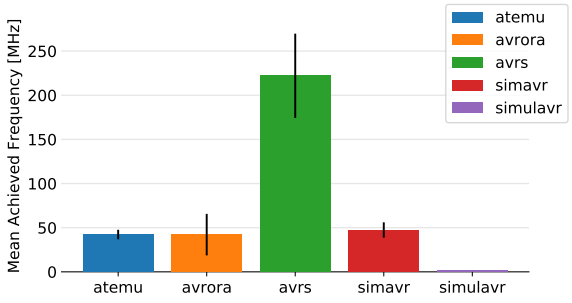


Figure 2: Achieved emulator frequencies across all benchmarks

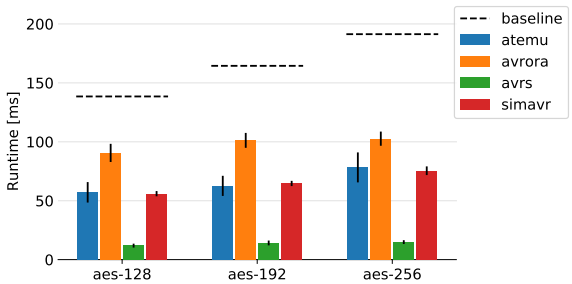


Figure 3: Performance of AES operations in emulators, slower than baseline omitted

Conclusions & Outlook

- Existing emulators lack completeness with respect to implemented instructions, even in mature, actively-developed emulators.
- Open-source toolchain support for ATtiny and ATxmega is error prone.
- AVRS can achieve competitive performance, as highlighted by figures 2 and 3, while also providing support for missing instructions.
- Fuzzing example shows ease of implementing analysis on top of AVRS.
- Use AVRS in future projects to evaluate symbolic execution/taint tracking on AVR devices, decode to IR allows exploring static analysis on AVR as well.

[1] Marius Muench, Dario Nisi, Aurélien Francillon, and Davide Balzarotti. Avatar2: A multi-target orchestration platform. In *Proc. Workshop Binary Anal. Res.(Colocated NDSS Symp.)*, volume 18, pages 1–11, 2018.
[2] Marius Muench, Jan Stijohann, Frank Kargl, Aurélien F rancillon, and Davide Balzarotti. What you corrupt is not what you crash: Challenges in fuzzing embedded devices. In *NDSS 2018, Network and Distributed Systems Security Symposium, 18-21 February 2018, San Diego, CA, USA*, S an Diego, UNITED STATES, 02 2018.
[3] Isaías González Pérez, Antonio José Calderón Godoy, Manuel Calderón Godoy, and Juan Félix González González. Survey about the utilization of open source arduino for control and measurement systems in advanced scenarios. application to smart micro-grid and its digital replica. In *ICINCO*, 2019.
[4] Ben L. Titzer, Daniel K. Lee, and Jens Palsberg. Avrora: Scalable sensor network simulation with precise timing. In *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks*, IPSN '05, Piscataway, NJ, USA, 2005. IEEE Press.
[5] Andrew Wickert, Chad Sandell, Bobby Schulz, and G.-H Ng. Open-source arduino-derived data loggers designed for field research. *Hydrology and Earth System Sciences Discussions*, pages 1–16, 12 2018.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Digital Twins for Cyber-Physical Threat Detection & Response

Laura Waltersdorfer, Matthias Eckhart, Andreas Ekelhart

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Cyber-Phyiscal Threats

- **Cyber-Physical Production Systems** (CPPS) are increasingly targeted by tailored threats
- **Malware attacks** can hinder operation, incur safety dangers and cause significant financial damages
- Sophisticated attacks (cf. TRITON [1]) are aimed at **Industrial Control Systems** and show the aggressiveness of such attacks

Security Testing

- Security testing in CPPS is inherently difficult due to the following reasons:
  (i) high costs for **custom infrastructure of testbeds**
  (ii) simulation of systems **highly complex**
  iii) **space constraints**
- Past attempts to perform penetration testing on CPPSs demonstrated **critical malfunctions**, uncontrolled disruption of operation and significant potential danger to human workers [2]

Detecting Cyber Threats with Digital Twins

Digital Twin Generation & Usage

- Digital Twins, virtual replicas of the CPPS, can be generated based on artefacts and engineering data (e.g. AutomationML files) in a cost-efficient way for security use cases [4]
- The real-system behavior is reflected in the digital twin, for example by passive state replication as shown in [3]
- The virtualisation enables the monitoring of the system according to process logic, assuring the adherence of specified states, comparing data and testing of future adaptations in a virtual, realistic environment

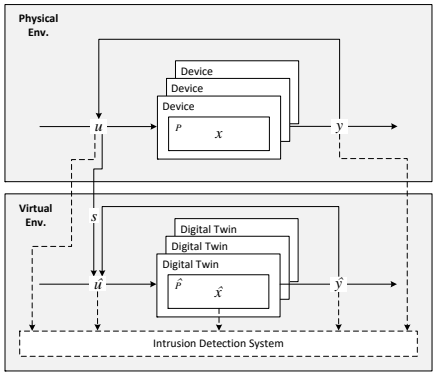


Figure 1: State Replication [3]

Advanced Security Use Cases for Digital Twins

Intrusion Detection

- The multi-dimensional nature of production processes must be addressed: process level (i.e. physics level), sensing and manipulation (e.g. PLC logic, sensor data and network level)
- Anomalies can be detected based on the specified behavior and provided safety rules

Response & Reconfiguration

- Proactive and reactive responses can be incorporated in the Digital Twin to increase resilience of CPPS

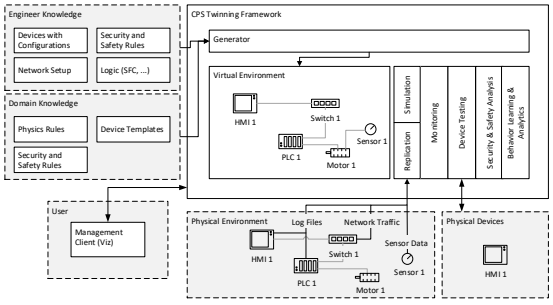


Figure 2: Digital Twin Framework Architecture

Conclusion

- Automatic generation of Digital Twins based on artefacts and engineering data is a cost-effective and low-effort measure to create an **enhanced security testing environment** for CPPS
- Digital Twins can help to increase the **security testing capabilities** under otherwise challenging circumstances and enable experts to play out different scenarios in the production system domain **without disturbing operation**

Future Work

- Behavior specification-based Intrusion Detection Systems seem promising for industrial systems to detect threats
- Integration of semi-automated deployment of reconfiguration **can increase the resilience of CPPS**

[1] AC Alessandro Di Pinto, Younes Dragoni, and Andrea Carcano. Triton: The first ics cyber attack on safety instrument systems. In Proc. Black Hat USA, pages 1–26, 2018.
[2] David Duggan, Michael Berg, John Dillinger, and Jason Stamp. Penetration testing of industrial control systems. Sandia national laboratories, 2005.
[3] Matthias Eckhart and Andreas Ekelhart. A specification-based state replication approach for digital twins. In Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and Privacy, pages 36–47, 2018.
[4] Matthias Eckhart and Andreas Ekelhart. Towards security-aware virtual environments for digital twins. In Proceedings of the 4th ACM workshop on cyber-physical system security, pages 61–72, 2018.

This research was further funded by the FFG under the industrial PhD program (grant no. 874644). Moreover, the financial support by the Christian Doppler Research Association, the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

Echoes of the Past: Recovering Blockchain Metrics From Merged Mining

Nicholas Stifter, Philipp Schindler, Aljosha Judmayer, Alexei Zamyatin, Andreas Kern, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research
TU WIEN
Christian Doppler Forschungsgesellschaft

Motivation

So far, the topic of *merged mining* has mainly been considered in a security context, covering issues such as mining power centralization or cross-chain attack scenarios. In this work we show that key information for determining blockchain metrics such as the *fork rate* can be recovered from data that is embedded in the blockchains of merge-mined cryptocurrencies.

Merged Mining as a Side Channel: The process of merged mining requires that miners must include block header candidates of a *parent* cryptocurrency which they are repurposing as part of the proof of work for a *child* cryptocurrency. This means that data from parent cryptocurrencies is embedded into the publicly available blockchains of merge-mined child cryptocurrencies. Bitcoin and Litecoin are the most prominent examples of parent cryptocurrencies for various merge-mined coins. We are the first to outline and consider this interesting source of blockchain data.

Validation and Improvement of Previous Measurements: Within this work, we analyze and evaluate the ability to use data from merge-mined cryptocurrencies to recover information about key metrics of parent cryptocurrencies. Specifically, we consider a reconstruction of the stale block rate and fork rate within Bitcoin and Litecoin based on merged mining data, and compare it to publicly available information derived from live measurements.

Public Source of Blockchain Data: The presented technique outlines how blockchain data that previously could only be obtained through active monitoring of the peer-to-peer network can be extracted from merge-mined cryptocurrencies. This new data source is particularly useful, as it also covers events from the past where no live monitoring data is available.

Methodology

- Categorization of different recoverable block types (see Figure 1)
- Analysis of data from nine merge-mined cryptocurrencies
- Comparison of recovered Bitcoin stale block rate to previous findings from Decker and Wattenhofer [1] (see Figure 2)
- Expanded analysis using multiple live monitoring data sources for comparison (see Figure 3)

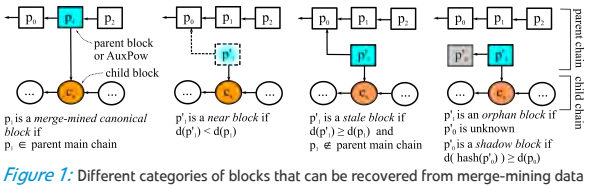


Figure 1: Different categories of blocks that can be recovered from merge-mining data

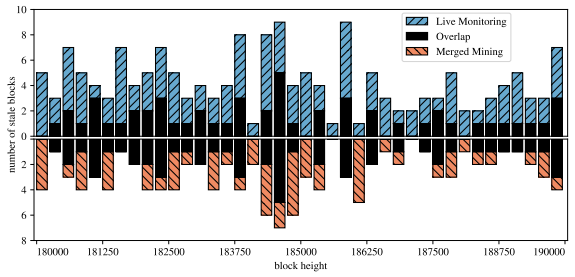


Figure 2: Comparison of stale block rates from the first live monitoring interval used in [1] and data recovered by our new technique. The overlap represents stale blocks present in both data sets.

Analysis Results

- Technique can be used to recover stale blocks and forks in Bitcoin and Litecoin over a long timespan
- Many stale blocks are uniquely identified through our technique, allowing more accurate fork rate estimates if all data is combined

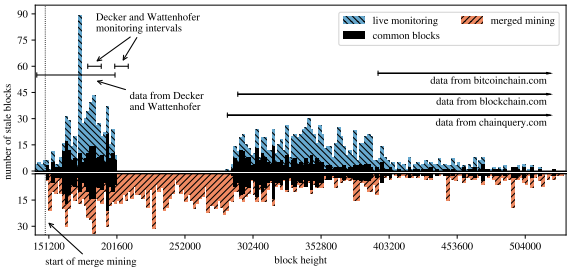


Figure 3: Comparison of stale block rates recovered by merged mining to combined live monitoring data from multiple public sources. A gap exists in live monitoring where no publicly available data sources could be located

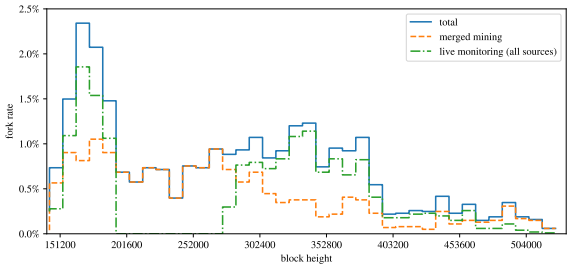


Figure 4: Updated estimate of the Bitcoin fork rate based on paper insights

Conclusion

- We outline that **merged mining provides a side channel for additional data** about the involved *parent* cryptocurrencies.
- We show that **merged mining data can be used to recover stale blocks and forks in the parent chain**, and may enable inference of other key metrics.
- Our analysis reveals a **sizable portion of forks and stale blocks that were not recognized through current live monitoring** activities, suggesting that this new approach serves as an important **complementary mechanism for correctly determining the fork rate**.
- We demonstrate that our **approach can be applied to other merge-mined parent cryptocurrencies** such as Litecoin.

[1] Christian Decker and Roger Wattenhofer. Information propagation in the bitcoin network. In *IEEE Thirteenth International Conference on Peer-to-Peer Computing (P2P)*, pages 1–10. IEEE, September 2013.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins

Matthias Eckhart, Andreas Ekelhart, and Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Background & Motivation

Cyber Situational Awareness refers to what an individual is aware of regarding events (e.g., issues, attack attempts) that occur in the cyber domain (e.g., networks) [6].

This state can be achieved at different levels [5]:

- Perception,
- Comprehension, and
- Projection of a situation.

Necessary for proper incident handling (e.g., understanding the anatomy of cyber attacks).

Problems

- Learning about incidents may be limited to the use of passive data collection approaches (active techniques may negatively affect the real-time performance).
- CPSs have stringent availability requirements (cannot be simply put out of operation for inspection).
- Difficult to retain valuable information for analysis.

Need: Ability to passively observe the state of CPSs without negatively affecting their operation.

A Cyber Situational Awareness Framework Based on Digital Twins

Contribution [4]: Improves cyber situational awareness by means of visualization and replaying recorded states from real devices to digital twins to reproduce events.

Overview

- Digital twins refer to simulated or emulated devices (e.g., PLCs, HMIs, sensors) that are connected to an emulated network [1, 3].
- Digital twins can be automatically generated from the specification of the CPS [2].
- We extend the open-source framework CPS Twinning: `https://github.com/sbaresearch/cps-twinning`
- Program states are passively mirrored from real devices to digital twins.
- Program and network layer of digital twins can be easily monitored.

Visualization

- Framework provides visualization panel that illustrates the CPS's topology.
- The digital twins' program variables can be monitored in real-time.
- Planned: visual feedback for user-defined alarms, detected intrusions and security metrics.

Record & Replay

- State replication [1] causes stimuli to be directly streamed to the digital twins (limits analysis capabilities).
- Framework stores stimuli for the purpose of replicating them at a later time.
- Users can step back or move forward in the state timeline of digital twins.
- Allows to easily inspect past behavior and establishes a reproducible analysis process.

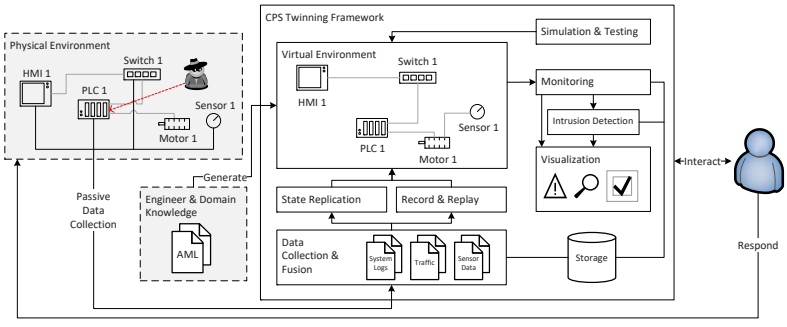


Figure 1: The architecture of the proposed digital-twin cyber situational awareness framework.

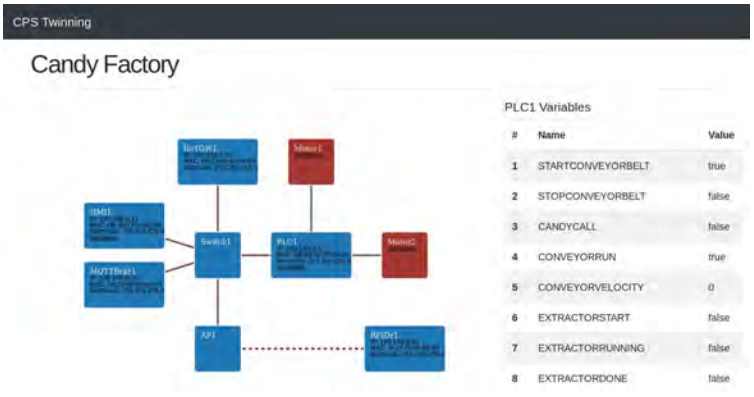


Figure 2: The visualization of digital twins, depicting the CPS's topology and program variables of (virtual) devices.

Use Cases

- **Risk Assessment:** Simulate threat scenarios and assess their impact.
- **Monitoring:** Indirectly observe the system behavior via digital twins.
- **Incident Handling:** Elucidate cyber incidents by means of visual analytics and the record-and-replay feature.

Conclusion

- Provides advanced monitoring, inspection, and testing capabilities.
- Based on the digital-twin framework CPS Twinning [2, 1].
- Further development required for improving visualizations and completing the record-and-replay feature.

[1] M. Eckhart and A. Ekelhart. A specification-based state replication approach for digital twins. In *Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy*, CPS-SPC '18, pages 36–47, New York, NY, USA, 2018. ACM.
[2] M. Eckhart and A. Ekelhart. Towards security-aware virtual environments for digital twins. In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security*, CPSS '18, pages 61–72, New York, NY, USA, 2018. ACM.
[3] M. Eckhart and A. Ekelhart. Digital twins for cyber-physical systems security: State of the art and outlook. In Stefan Biffl, Matthias Eckhart, Arndt Lüder, and Edgar Weippl, editors, *Security and Quality in Cyber-Physical Systems Engineering*, volume 1. Springer International Publishing, 2019.
[4] M. Eckhart, A. Ekelhart, and E. Weippl. Enhancing cyber situational awareness for cyber-physical systems through digital twins. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1222–1225, Sep. 2019.
[5] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1):32–64, 1995.
[6] U. Franke and J. Brynielsson. Cyber situational awareness – a systematic review of the literature. *Computers & Security*, 46:18 – 31, 2014.

This research was further funded by the FFG under the industrial PhD program (grant no. 874644). Moreover, the financial support by the Christian Doppler Research Association, the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

Event-Based Failure Prediction in Distributed Business Processes

Michael Borkowski, Walid Fdhila,
Matteo Nardelli, Stefanie Rinderle-Ma, Stefan Schulte

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

Problem & Motivation

- How can contextual data of external EBS be combined with process-intrinsic events of a (distributed) business process in order to predict process failures?
- How to formalize and automate failure prediction in (distributed) process settings?
- Is the approach feasible when used in a real-world scenario, and how well does the failure prediction perform?
- What is the impact of the distribution of processes on failure prediction?

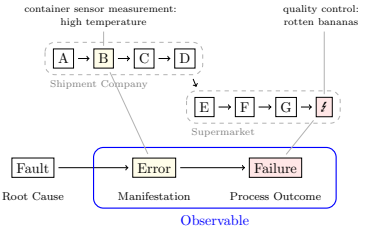


Figure 1: Example of Faults, Errors and Failures as Parts of a Process

Machine Learning Failure Prediction

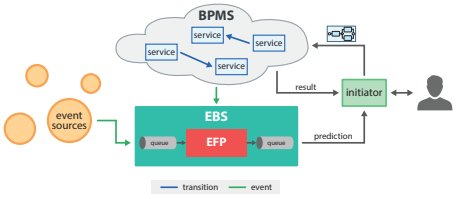


Figure 2: System Architecture

- Initiator
  - starts the business process execution and the EFP component.
  - forwards process execution results and failure predictions to users.
  - can also be distributed.
- Event-based Failure Prediction component (EFP)
  - analyses intrinsic and contextual events and predicts failures.
  - includes an online training subsystem to build an artificial neural network model for failure. prediction

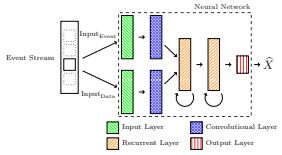


Figure 3: Schematic Representation of the Proposed ANN Model

- Event-based System (EBS)
  - creates and manages EFP instances.
  - implements a message queue system.

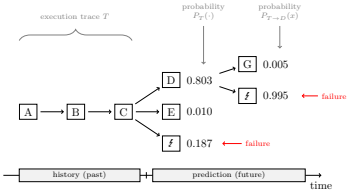


Figure 4: Trace Tree Example

Data Sets

- Real world data set (from BPI Challenge 2017)
- Synthetic data set
  - uses realistic process collaboration (supply chain).
  - simulated using the Cloud Process Execution Engine (CPEE).
  - events stored in an Extensible Event Stream (XES) file.
- Fault Injection
  - step-indicated faults.
  - event-indicated faults.
  - data-indicated faults.

Analysis

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1) \qquad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \quad (3)$$

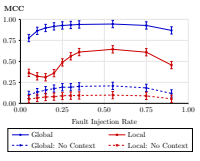


Figure 5: Classifier MCC Using Synthetic Data Set, over Varying Fault Rates

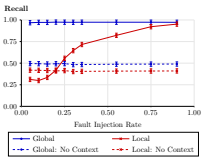


Figure 7: Classifier Recall Using Synthetic Data Set, over Varying Fault Rates

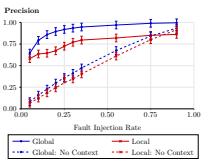


Figure 6: Classifier Precision of Synthetic Data Set, over Varying Fault Rates

Metric	Mean	σ
Precision	0.873	0.344
Recall	0.971	0.307
MCC	0.879	0.331

Table 1: Key Metrics for Evaluation Using Real-World Data Set (Mean Values for 10-fold Cross Validation)

Conclusion

- ANN-based failure prediction for DBPMS combined with EBS
- Satisfactory results with high precision, recall, and MCC
- Data availability impacts significantly the classification performance
- Increasing fault rate leads to an increase in both precision and recall

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
Ein Fonds der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Exploiting ICMPv6 Error Messages for Reconnaissance

Florian Holzbauer, Markus Maier, Johanna Ullrich

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

- Reconnaissance in IPv6 remains an open problem due to its sheer address space.
- However, the amount of error messages usually exceeds the amount of positive replies in IPv6.
- We investigate whether error message allow to infer the deployment status of an IPv6 network.

Methodology

- Behaviors of virtual router appliances are monitored in a lab setup.
- Results are cross-checked with response behavior in the wild.
- Contribution of error messages to active network detection is shown.

Response Behavior in the Wild

Based on a list of addresses known to be active, we generated test cases that represent (I) probing of an active network, and (II) probing of an inactive network.

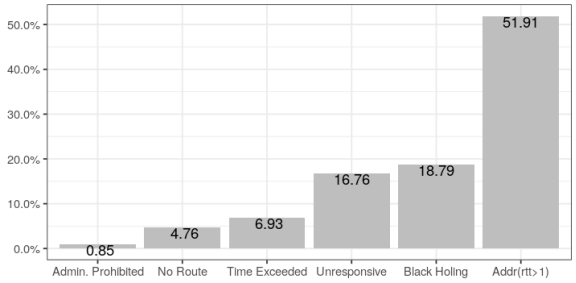


Figure 1: Active network responses

More than 50% of active networks respond with "Address Unreachable" with a RTT > 1s. In contrast, inactive networks react differently. We found requests to 20.8% of inactive networks result in a routing loop.

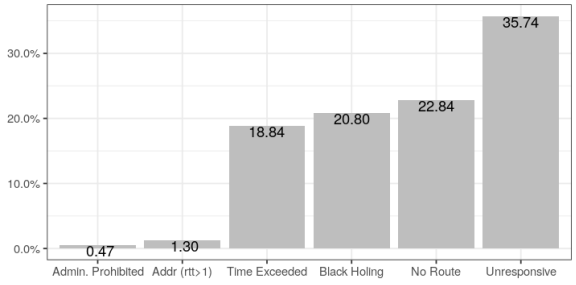


Figure 2: Inactive network responses

Conclusion

- Error messages can be used to infer the deployment status of an IPv6 network. Therefore we recommend a two-step approach for IPv6 measurements: 1) Detect active subnets. 2) Investigate active subnets for active hosts.
- We detected numerous routing loops in IPv6.

Lab Environment

We investigated response behavior of routers in a laboratory setup and found error message types to be ambiguous. For example, "Address Unreachable" is used for an inactive host in an active network, but also in case of an active reject route.

	Protocols	Act. NW, Act. Host	Act. NW, Inact. Host	Inact. NW, Inact. Host	Act. ACL	Act. Rej. Route	Rout. to Inc. Interface
Test Case No.		1	2	3	4	5	6
Cisco IOS 15.2.4	All	Reply	Address Unreachable	No Route	Admin. Prohibited	Reject Route	Time Exceeded
Cisco CRS1000V	All	Reply	Address Unreachable	No Route	Admin. Prohibited	Reject Route	-
Juniper VMx 17.1	All	Reply	Address Unreachable	No Route	Admin. Prohibited	Address Unreachable	Time Exceeded
HPE VSR1000	All	Reply	-	-	-	-	Time Exceeded
Mikrotik 6.45	All	Reply	Address Unreachable	No Route	No Route	No Route	Time Exceeded
OpenWRT 19.07	ICMP	Reply	Address Unreachable	Failed Policy	Port Unreachable	Failed Policy	Failed Policy
	TCP/UDP	Reply	Reply	Failed Policy	Reply	Failed Policy	Failed Policy

Table 1: Recorded Message Originating Behavior of Router Operating Systems

Response Timings in the Wild

There is a difference in timing of "Address Unreachable" messages that allows to gain insight into the remote network's status.

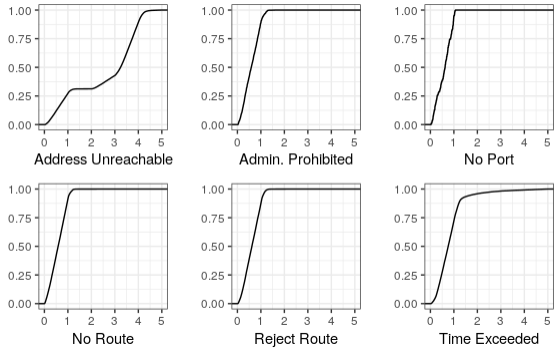


Figure 3: CDF of round-trip times (in seconds) per message type

Demonstration

- *764 Enterprise Networks*: We scanned all directly allocated /48 networks from Gasser's hitlist. Based on Addr (rtt>1) we found around 39.5 active subnets for each network, while the hitlist only shows 3.23 subnets per network.
- *1 ISP Network*: Error Messages allowed us to differentiate between custom home routers and ISP proprietary routers in use. We detected only 0.5% of home connections use custom routers.
- *1 ISP Business Network*: We know from at least one network range of our research partner to be active. While we detected 41 other active networks through Echo Replies, 29 were solely detected by Addr(rtt>1), enhancing the results by about 70%.

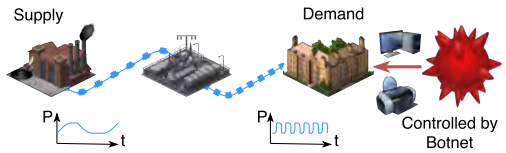
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Grid Shock:
Coordinated Load-Changing Attacks on Power Grids

Adrian Dabrowski, Johanna Ullrich, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

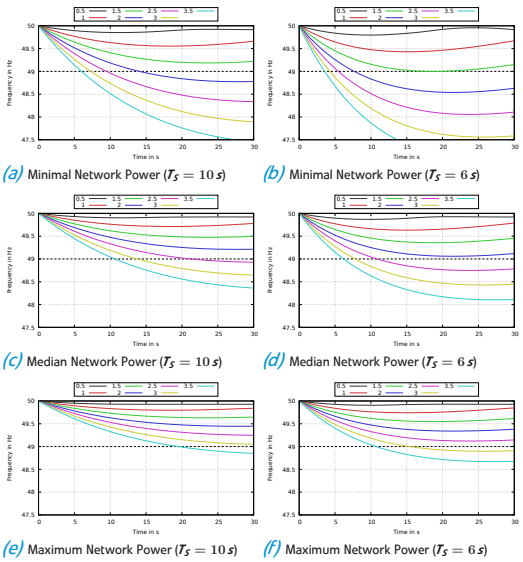
Problem & Motivation



Electric power grids are critical infrastructure. For reliable operation, providers have to continuously maintain a balance between supply and demand to keep the grid's nominal frequency of 50 Hz. In our work, we assume an adversary aiming to destabilize the power grid. Therefore, she builds a botnet of zombie computers and modulates their power consumption in a concerted fashion.

Static Load Attacks

In static load attacks, the adversary synchronously increases the electric load of the bots. The impact on the frequency is shown for a grid with high rotational inertia ($T_S = 10\,s$), i.e., predominantly fed by conventional power plants, and low rotational inertia ($T_S = 6\,s$), i.e., fed by a high share of renewables, at different levels of total network power. Static load attacks are in multiples of the ENTSO-E reference incident (3,000 MW).



(a) Minimal Network Power ($T_S = 10\,s$) (b) Minimal Network Power ($T_S = 6\,s$)

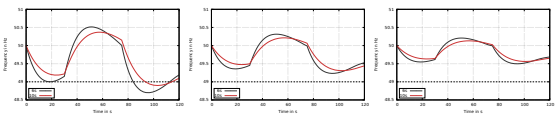
(c) Median Network Power ($T_S = 10\,s$) (d) Median Network Power ($T_S = 6\,s$)

(e) Maximum Network Power ($T_S = 10\,s$) (f) Maximum Network Power ($T_S = 6\,s$)

Figure 1: Impact of Static Load Attacks on Grid Frequency

Dynamic Load Attacks

In dynamic load attacks, the adversary increases the load to the maximum and waits for the primary control to be activated; then, she decreases the load deactivating primary control again.



(a) Minimal Network Power (b) Median Network Power (c) Maximum Network Power

Figure 2: Impact of Dynamic Load Attacks on Grid Frequency

Controllable Load

From our own measurements and data sheets, we compiled a table of controllable load by PC components and Internet-of-Things devices encompassing the potential for increasing and/or decreasing power, latencies of power modulation, and the amount of controllable load.

		Pwr Control		Latency		
Device	Type	Inc.	Dec.	on	off	Δ Load
CPU	Core2 Duo		○	20-60 ms	20-60 ms	35 W
	i3		○	20-60 ms	20-60 ms	55-73 W
	i5		○	20-60 ms	20-60 ms	73-95 W
	i7		○	20-60 ms	20-60 ms	77-95 W
	i7-E		○	20-60 ms	20-60 ms	130-150 W
GPU	Low-end		○	20-60 ms	20-60 ms	20-76 W
	Mid-end		○	20-60 ms	20-60 ms	102-151 W
	High-end		○	20-60 ms	20-60 ms	150-238 W
	Top-end		○	20-60 ms	20-60 ms	201-297 W
HDD			○	20-60 ms	20-60 ms	3-7 W
Screen TFT	size dep.			1-5 s	5-10 s	60-100 W
Laser Printer	SOHO		○	1-3 s	5-10 s	800-1300 W
Smart Air Cond.			○	1-10 s		600-1000 W
Smart Thermostat	elec. Heating		○	1-10 s		1-15 kW
Smart Oven			○	1-10 s		2-3 kW
Smart Refrigerator			○	1-10 s		300-500 W
Smart Kettle			○	1-10 s		1000-1500 W

Table 1: Latency and Achievable Load Differences

Conclusion

An adversary does not have to rely on smart grid features to modulate power consumption, given that an adequate communication infrastructure for striking the (legacy) power grid is currently nearly omnipresent: the Internet, to whom more and more power-consuming devices are connected. **Our simulations estimate that between 2.5 and 9.8 million infections are sufficient to attack the European synchronous grid.**

A. Dabrowski, J. Ullrich, E. Weippl, 2017, *Grid Shock: Coordinated Load-Changing Attacks on Power Grids*, Annual Computer Security Applications Conference (ACSAC).

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

HydRand: Efficient Continuous Distributed Randomness*

Philipp Schindler, Aljosha Judmayer, Nicholas Stifter, and Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Introduction to Distributed Randomness

A reliable source of randomness is not only an essential building block in various cryptographic, security, and distributed systems protocols, but also plays an integral part in the design of many new blockchain proposals.

Distributed Randomness as Alternative for Proof of Work: Lately, randomness beacons have received increased attention, in part because generating shared randomness is proving to be a vital component of most distributed ledger approaches that aim to replace the computationally intensive Proof-of-Work (PoW) mechanism. Specifically, Proof-of-Stake (PoS) blockchain proposals, which rely on virtual resources in the form of digital assets, call for manipulation-resistant and unpredictable leader election as part of a secure protocol design. The distributed generation of trustworthy random values can hence be considered a complementary problem to the development of such protocols.

Our Contribution: We improve upon previous random beacon approaches with HydRand, a novel distributed protocol based on publicly verifiable secret sharing (PVSS) to ensure unpredictability, bias-resistance, and public-verifiability of a continuous sequence of random beacon values. Furthermore, HydRand provides **guaranteed output delivery of randomness at regular and predictable intervals** in the presence of adversarial behavior and does not rely on a trusted dealer for the initial setup. Compared to existing PVSS-based approaches that strive to achieve similar properties, our solution improves scalability by lowering the communication complexity from $O(n^3)$ to $O(n^2)$. Furthermore, we are the first to present a **detailed comparison** of recently described schemes and protocols that can be used for implementing random beacons.

Key Properties

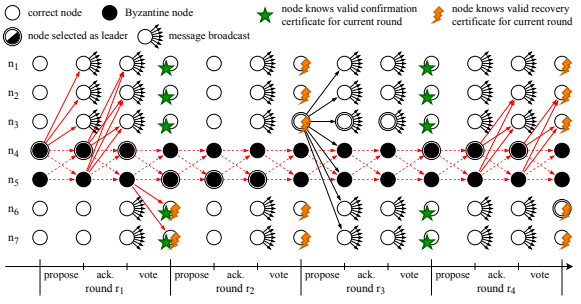
- Availability/Liveness
- Unpredictability
- Bias-Resistance
- Public-Verifiability
- Guaranteed Output Delivery

Application Areas

- Blockchains and DLTs
- Smart Contracts
- Generation of Protocol Parameters
- Privacy-Preserving Messaging
- Anonymous Browsing
- E-Voting Protocols
- Publicly Auditable Sections
- Gambling and Lottery Services

Our Random Beacon Protocol: HydRand

- Stand-alone Protocol Design
- Open Source Implementation Available on Github
- Easy Setup: No Trusted Dealer and No DKG Required
- Strong Guarantees while Improving Performance and Scalability



Comparison

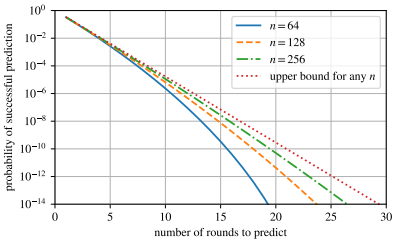
Recent years have seen a substantial amount of new research in this field. We aid the selection process for different use cases by presenting the first comprehensive comparison of available protocol designs.

	Communication model	Liveness / failure probability$^\diamond$	Comm. complexity (overall protocol)	Unpredictability	Bias-Resistance	Comp. complexity (per node)	Verif. complexity (per passive verifier)	Characteristic cryptographic primitive(s)	Trusted dealer or DKG required
Algorand	semi-syn.	10^{-12}	$\mathcal{O}(cn)^*$	↯↗	✗	$\mathcal{O}(c)^*$	$\mathcal{O}(1)^*$	VRF	no
Cachin et al.	asyn.	✓	$\mathcal{O}(n^2)$	✓	✓	$\mathcal{O}(n)$	$\mathcal{O}(1)$	uniq. thr. sig.	yes
Caucus	syn.	✓	$\mathcal{O}(n)$	↯↗	✗	$\mathcal{O}(1)$	$\mathcal{O}(1)$	hash func.	no
Dfinity	syn.	10^{-12}	$\mathcal{O}(cn)$	✓	✓	$\mathcal{O}(c)$	$\mathcal{O}(1)$	BLS sig.	yes$^\#$
Ouroboros	syn.	✓	$\mathcal{O}(n^3)$	✓	✓	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	PVSS	no
Ourob. Praos	semi-syn.	✓	$\mathcal{O}(n)^*$	↯↗	✗	$\mathcal{O}(1)^*$	$\mathcal{O}(1)^*$	VRF	no
Proof-of-Work	syn.	✓	$\mathcal{O}(n)$	↯↗	✗	very high§	$\mathcal{O}(1)$	hash func.	no
Proof-of-Delay	syn.	✓	$\mathcal{O}(n)$	✓	✓	very high§	$\mathcal{O}(\log \Delta)^\circ$	hash func.	no
RandShare	asyn.	✗†	$\mathcal{O}(n^3)$	✓	✓	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	PVSS	no
RandHound	syn.	0.08%	$\mathcal{O}(c^2n)^\ddagger$	✓	✗	$\mathcal{O}(c^2n)$	$\mathcal{O}(c^2n)$	PVSS	no
RandHerd	syn.	0.08%	$\mathcal{O}(c^2 \log n)^\ddagger$	✓	✓	$\mathcal{O}(c^2 \log n)$	$\mathcal{O}(1)$	PVSS/CoSi	yes$^\#$
Scrape	syn.	✓	$\mathcal{O}(n^3)$	✓	✓	$\mathcal{O}(n^2)$	$\mathcal{O}(n^2)$	PVSS	no
HydRand	syn.	✓	$\mathcal{O}(n^2)$	↯✓	✓	$\mathcal{O}(n)$	$\mathcal{O}(n)$	PVSS	no

Evaluation

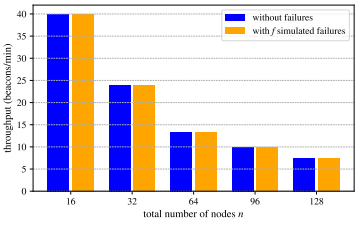
Our security analysis shows that HydRand achieves the desirable properties of a random beacon protocol: liveness, guaranteed output delivery, unpredictability, bias-resistance, and public-verifiability.

- Unpredictability guarantees for different numbers of nodes, assuming a 33% adversary

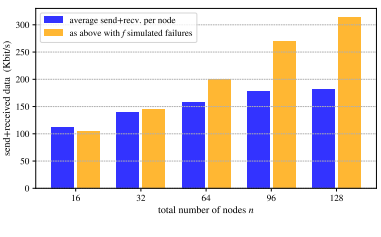


We performed large-scale performance tests by executing the HydRand protocol on up to 128 Amazon EC2 t2.micro instances (1 GiB of RAM, one virtual CPU core, 60-80 Mbit/s network bandwidth). Instances were spread equally over multiple data centers in eight AWS regions. The results highlight that the presented HydRand protocol is feasible for realistic deployment scenarios.

- Number of Random Beacon generated per minute, with and without simulated failures



- Average used bandwidth per node in Kbit/s, with and without simulated failures



* To appear in the Proceedings of the IEEE Symposium on Security and Privacy (IEEE S&P) 2020.
A short introduction video is available at https://www.youtube.com/watch?v=0ZrN6fmM4TQ.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschungsförderungsgesellschaft
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Implanting Security Features into Resource Constrained Embedded Systems

Matthias Wenzl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Background

IoT, IIoT, & Embedded Systems

In the last decade, emerging use cases like "Ambient Assistive Technologies", "Car2X" communication, "Smart Homes", "Smart Cities" and Industry 4.0 transformed computer systems to ubiquitous companions in our daily lives. The majority of these use cases consists of embedded computing devices that are highly specialized for their particular purpose.

Requirements

- *Physical:* Operation in harsh conditions, size (e.g. intelligent drilling head).
- *Energy:* Long mission times (e.g. wild-life tracking).
- *Memory:* At least X% of volatile and non-volatile memory must be free during deployment.
- *Real-time requirements:* Adherence to computation time deadlines during operation (e.g., parking assistant).

Hardware Heterogenity

- *ATTiny85:* 8-bit ISA, few kb RAM, few kb Flash, no MMU/MPU, low clock frequency, no pseudo random number generator in hardware
- *Jetson TX2:* 64-bit ISA, lots of GB RAM, few GB Flash, complete MMU, > 1 GHz clock frequency, dedicated hardware acceleration units

Software Diversity

- *General purpose OS's:* Windows 10 IoT Core, Linux, *BSD
- *Embedded OS's:* RIOT, TI-RTOS, Free-RTOS, Contiki, QNX
- *Standalone applications:* no operating system at all

Problem & Motivation

IoT/IIoT Security Incidents

- S7 Simatic SPS (2019) Upload of an arbitrary program by utilizing protocol vulnerability
- BMW (2018) Remote exploitable vulnerabilities using the infotainment
- University IoT (2017) Take over of networked luminance sensors, lights and vending machines
- POS (2017) Take over of restaurant receipt printers using remote execution flaws
- Mirai (2016 and before) Botnet mostly exploiting weak default credentials of certain IoT devices
- Jeep (2015) Remote code execution in the Infotainment system

Why are Embedded System so vulnerable?

Figure 1: Availability of assorted security features in embedded systems without standard operating system. ("Ghost in the Machine: Challenges in Embedded Binary Security", Wetzels, Usenix Enigma, 2017)

An easy solution?

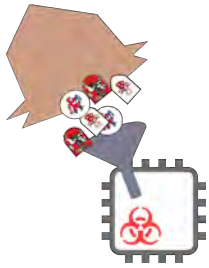


Figure 2: Application domain specific requirements as well as HW and SW diversity hinder the application of common security features in firmware.

Proposed Solution

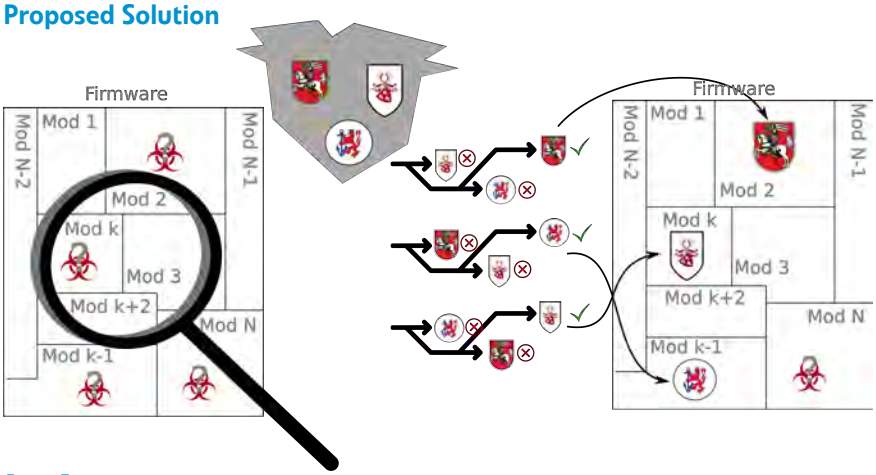


Figure 3: Detection of vulnerable points → Select security feature → Implant security feature → Perform global optimization in order to maximize the number of protected vulnerable points with respect to the system's constraints.

- Detection of potential vulnerable spots → taint tracing
- Select appropriate security feature with respect to: run-time, memory complexity, effectiveness.
- Implant security feature using binary rewriting.
- Perform global optimization in order to maximize the number of protected vulnerable points with respect to the system's constraints.
  - **Strategy 1:** Maximize the number of protected vulnerable spots, but allow some to be protected by less effective mitigation techniques
  - **Strategy 2:** Protect the vulnerable spots with high ratings with the most efficient exploit mitigation techniques available, but have less protected in total.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.
This work is partly supported by the Austrian Forschungsförderungsgesellschaft (FFG) in the context of project ISaFe (871230).

Mobile Atlas
A Scalable Way to Measure Cellular Networks

Wilfried Mayer, Adrian Dabrowski, Gabriel Gegenhuber

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

- **Mobile data traffic is surging**, data roaming is increasingly popular.
- Large-scale internet measurement platforms only exist for fixed line connections, **mobile networks mostly ignored**.
- Cellular networks **differ in terms of measurement requirenments** and objectives (calling, messaging, metered connections, zero rating offers, etc).
- A mobile-app-based approach impairs background user activity and is not suited for accurate technical measurements.
- It is especially hard to measure the roaming situation in a **controlled environment**.

Geographic Decoupling of Modem and SIM

Physically moving devices and SIM cards between countries to enable measurements in a romaing environment is costly and does not scale well.

Therefore we developed an approach to **geographically detach the SIM card** from the modem by **tunneling the SIM card's protocol** over the public Internet and emulating its signal on the LTE modem. This allows us to test roaming effects on a large number of operators without physically move any hardware between different countries.

Measurement Framework Architecture

- The **measurement probe** consists of a single board computer with an LTE modem that is connected via USB. The board' GPIO pins are connected to the SIM interface on the modem and emulate a physical SIM card.
- The **SIM provider** consists of multiple SIM cards with individual SIM card readers that are connected via USB to a host computer. Since those SIM card readers are cheap and easily available, our framework allows us to add arbitrary distributed SIM cards from other sources and test them throughout our measurement probes.

The distinct parts of our infrastructure **communicate through a VPN**. The **tunnel connection** between a measurement probe and the SIM provider **forms a virtual circuit**, where one SIM card can be connected to exactly one modem. Being able to connect any SIM card with any radio module regardless of geography **boosts automatability** of tests across a large number of SIMs and radio networks.

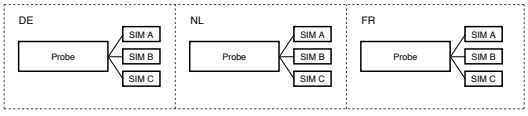


Figure 1: Naive architecture, exponential growth of used SIM cards, hard to maintain

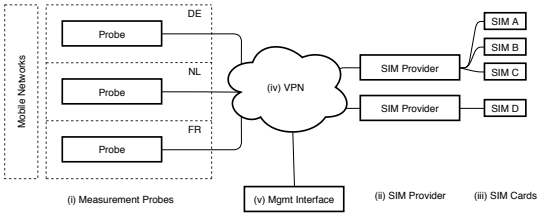


Figure 2: Our architecture allows every probe to use SIM cards attached to SIM providers, independent from the geographical location of probe and SIM card

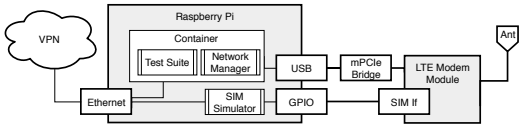


Figure 3: Design of our measurement probes

Current Deployment State

- Deployed in Europe
  - Austria, France, Germany, Netherlands, Spain, United Kingdom
  - Total of 290 million people
- Open and scalable design
  - Easily extendable

Figure 4: Current coverage of our measurement framework

Accomplishments

- Our Framework **enables qualitative measurements** in a controlled environment.
- **Scalable, open design, easily extensible**
  - Linear increase in costs per new probe
  - Practically no additional costs for new SIM cards

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Pay-To-Win: Incentive Manipulation Attacks on Cryptocurrencies

Aljosha Judmayer, Nicholas Stifter, Alexei Zamyatin,
Itay Tsabary, Ittay Eyal, Peter Gaži, Sarah Meiklejohn, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

*"The system is secure as long as **honest** nodes collectively control more CPU power than any cooperating group of attacker nodes."*

Satoshi Nakamoto [4]

- Does this also hold for **rational** nodes?
- **Bribing attacks** target this rationality assumption by offering bribes to manipulate the incentives of miners.
- They have first been described by Bonneau in [1] and extended in several papers (see [2]).

All previous approaches face two main **counter arguments**:

1. Bribing attacks require an extensive amount of funds.
2. Miners executing or joining such attacks risk to harm their own income stream due to value loss.

We present **two attacks** which show that those assumptions do not necessarily hold.

Out-of-band tx Exclusion and Ordering Attack

- The novelty of our attacks is that they can be executed and funded on a different (funding-)cryptocurrency/blockchain.
- The payment of bribes happens **out of band**, therefore counter argument two cannot be applied.
- This attack is cheaper than previous attacks because:
  - It changes the reward model towards always rewarding complacent miners.
  - The goal of the attack is shifted towards transaction exclusion and ordering of transactions.

Main idea:

- Attacker publishes half-finished blocks (block templates) in a smart contract on a different funding cryptocurrency.
- The smart contract then rewards collaborating miners for providing a PoW for these block templates.
- Complacent miners can be sure to receive a reward ⇒ **no risk**.

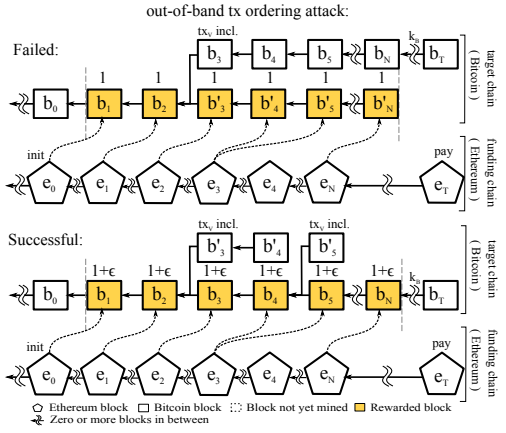


Figure 1: Block templates are published as transactions in the funding cryptocurrency and refer to blocks in the target cryptocurrency. The payouts are performed in block e_T. The colored blocks are rewarded by the contract, either only with their original value (reward + fee normalized to 1) or with an additional ϵ if the attack was successful.

Out-of-Band tx Revision Attack

- This attack incentivizes double-spend collusion, but also enables the attacker to exclude and order transactions.
- Due to the changed reward structure compared to previous attacks, it is cheaper than comparable attacks, i.e., the whale attack [3]. The operational costs for smart contract execution for a 6 block fork are around 30 USD.
- Additionally, crowdfunding mechanisms can be used to make the attack arbitrarily cheap in theory (see [2]).

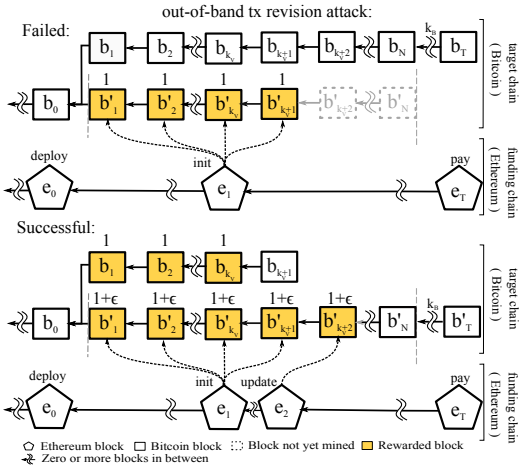


Figure 2: After k_V blocks on the target chain have passed, the attack contract is initialized with $k_V + 1$ block templates. The double-spend transaction(s) are included in block b'_1. The payouts are performed in block e_T. The colored blocks are rewarded by the attack contract, either with their original value (reward + free normalized to 1) or with an additional ϵ if the attack was successful.

ω	Whale costs	P2W costs c_{failed}	P2W costs $c_{success}$	N	P2W costs $c_{expected}$	N expect.
0.532	2.93e+23	6953	596	515	136	55
0.670	999.79	554	122	41	95	14
0.764	768.09	298	104	22	92	10
0.828	1265.14	216	98	16	90	9
0.887	1205.00	176	94	13	89	8
0.931	1806.67	149	92	11	89	8
0.968	2178.58	122	90	9	89	7
0.999	2598.64	108	90	8	89	7

Table 1: Comparison of attack costs given in BTC for the whale attack [3] and this attack with $k_V = 6$, $r_b = 13.5$, $c_{operational} = 0.5$ and $\epsilon = 1$.

Conclusion

- Incentive manipulation attacks provide a broader attack surface than previously anticipated.
- They can be executed from other (funding-)cryptocurrencies to attack a certain target cryptocurrency.
- They can be made cheaper, either through different attack goals (transaction exclusion), rewarding schemes (always pay) or through crowdfunding.

[1] Joseph Bonneau. Why buy when you can rent? bribery attacks on bitcoin consensus. In *BITCOIN'16: Proceedings of the 3rd Workshop on Bitcoin and Blockchain Research*, February 2016.
[2] Aljosha Judmayer, Nicholas Stifter, Alexei Zamyatin, Itay Tsabary, Ittay Eyal, Peter Gaži, Sarah Meiklejohn, and Edgar Weippl. Pay-to-win: Incentive attacks on proof-of-work cryptocurrencies. Cryptology ePrint Archive, Report 2019/775, 2019. https://eprint.iacr.org/2019/775.
[3] Kevin Liao and Jonathan Katz. Incentivizing blockchain forks via whale transactions. In *International Conference on Financial Cryptography and Data Security*, pages 264–279. Springer, 2017.
[4] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, Dec 2008. Accessed: 2015-07-01.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Power Grid Modelling: How Open Geoinformatic Data Benefits Security Research

Anja Klauzer, Johanna Ullrich

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Motivation

- **Detailed models** of the power grid are **inaccessible for the security community**. However, power lines, substations and plants appear to be well documented by OpenStreetMap.
- **Is open geodata a reliable source of information**? Does open geodata provide worthwhile insights for adversaries?
- In this work, we **model** the Austrian power grid based on **OpenStreetMap Data**.

Processing the data in GIS

- **Import data** for each state where **"power" is the key attribute** and "line", "minor line", and "substation" are the values.
  Clipping all vector layers, as the data are not linked to each other.
- **Generate geometric measurements** based on a selected CRS.
- **Calculate a buffer of 1km** for all substation objects to find out which lines might be connected to them.
- **Clipping all lines with the municipalities** where they run through.

Terms & Definition

Geographic Information System (GIS)
- Computer-aided system
- Model and process spatial data

Geodata
- Describe an object, either directly (by coordinates) or indirectly (e.g., by postal code), a landscape or by its position in space
- Can be linked to each other in order to create detailed queries and analyses

OpenStreetMap
- Built by a community of mappers that contribute and maintain open data about objects on the earth's surface

QGIS
- Open source GIS for viewing, editing, and capturing spatial data
- OSM plugin allows the access to up-to-date OSM data, and simple export to an easy-to-use Shapefile or SQLite database

Maps & Tables

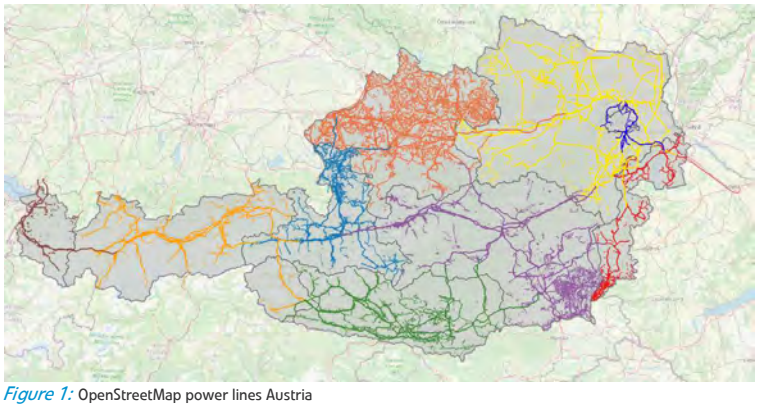


Figure 1: OpenStreetMap power lines Austria

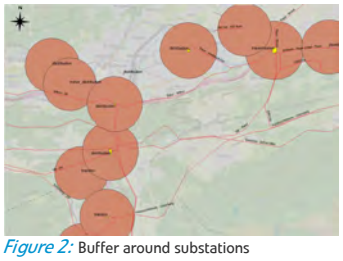


Figure 2: Buffer around substations

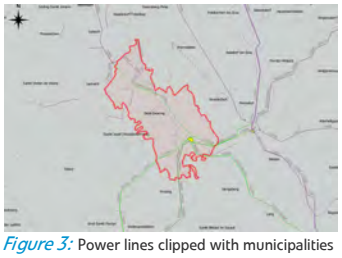


Figure 3: Power lines clipped with municipalities

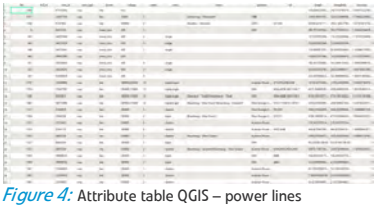


Figure 4: Attribute table QGIS – power lines

Figure 5: Exported attribute table – including coordinates

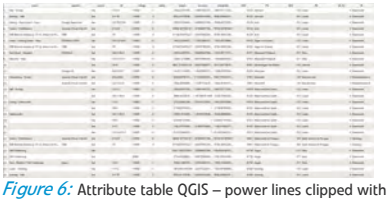


Figure 6: Attribute table QGIS – power lines clipped with municipalities

Conclusion & Future Work

- Based on the wide range of accessible geodata, we were **able to model the Austrian power grid**.
- Based on this model, we plan to **investigate attack strategies** using network analysis in the future.
- **Geodata provides** worthwhile **insights** for the security community and adversaries alike.
- Security-by-obscurity **cannot protect large-scale critical infrastructures** like the power grid.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien Ein Fonds der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee SCIENCE INNOVATION

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Proof-of-Blackouts?
How Proof-of-Work Cryptocurrencies Could Affect Power Grids

Johanna Ullrich, Nicholas Stifter, Aljosha Judmayer, Adrian Dabrowski, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Motivation

- Power grid operators have to maintain a balance between supply and demand. An imbalance causes a frequency shift from the set point of 50 Hz. While small deviations are normal, larger deviations cause emergency routines like load shedding.
- Proof-of-Work cryptocurrencies consume high loads of electric power from the grid. We investigate whether the power consumption of Bitcoin and Ethereum in Europe might destabilize the Synchronous Grid of Continental Europe spanning over 24 countries in Europe and Northern Africa.

Threat Model

An (occassional or malicious) incident leads to the outage of miners eventually causing totalled fluctuations on power consumptions.

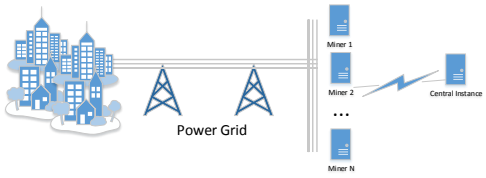
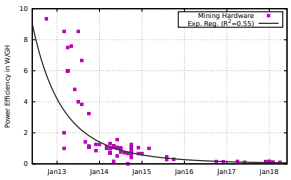
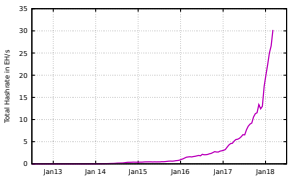


Figure 1: Threat Model

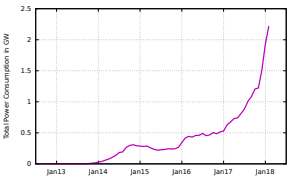
Results: Bitcoin



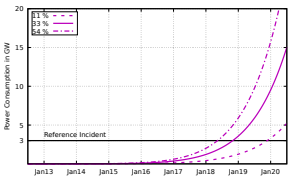
Power Efficiency of ASIC Miners



Total Hash Rate

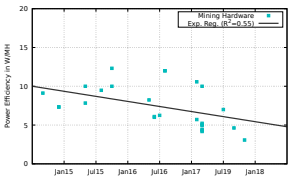


Total Power Consumption

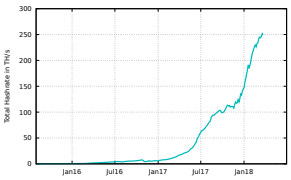


Projection of Consumption in Europe

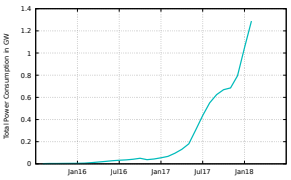
Results: Ethereum



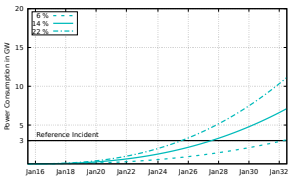
Power Efficiency of GPUs



Total Hash Rate



Total Power Consumption



Projection of Consumption in Europe

Methodology

1. Assessment of available power consumption models
2. Collection of data sheets for mining hardware
3. Regression analysis of mining hardware's power efficiency
4. Estimation of hash rate from public blockchains
5. Projection of total power consumption based on a mining mix
6. Attribution of power consumption to geographic areas
7. Comparison with reference values of respective power grids

Conclusion

- In Europe, Bitcoin and Ethereum draw power in the order of magnitude of a nuclear power plants.
- An incident leading to an outage of miners is able to destabilize the European power grid and might eventually lead to blackouts.
- In comparison to previous works focusing on botnets, such an incident decreases the grid's electric load – a situation which is more difficult to handle by grid operators.

J. Ullrich, N. Stifter, A. Judmayer, A. Dabrowski, E. Weippl, 2018, *Proof-of-Blackouts? How Proof-of-Work Cryptocurrencies Could Affect Power Grids*, International Symposium on Research in Attack, Intrusion, and Defenses (RAID).

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Securing the Testing Process for Industrial Automation Software

Matthias Eckhart, Andreas Ekelhart

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

The software testing process represents an **attractive attack target**:

- Risk of software piracy & theft of IP
- Covert attacks based on know-how gained via stolen artifacts (cf. Stuxnet)
- Means to conceal injected malicious code
- Potential damages to physical systems during test execution

Conducting security analyses (e.g., as per the VDI/VDE 2182 [7] guideline) of the testing process is **challenging**:

- Requires expert security know-how
- Is complex and effortful to perform
- Insufficient tool support available

Need: Framework to (semi-)automate security risk assessments with flexible assessment scope

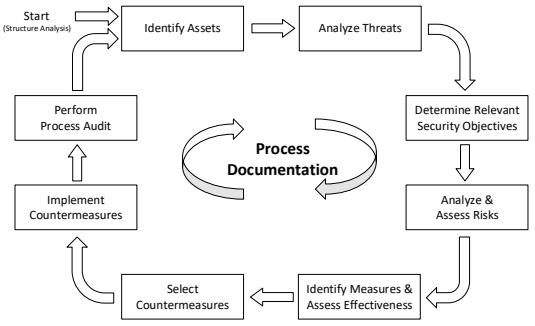


Figure 1: The procedural method according to the VDI/VDE 2182 [7] guideline.

Semi-Automated Security Analysis Framework

Contribution [1]: Provides the capabilities to conduct security analyses of an organization's software testing process for industrial automation software in a semi-automated manner.

Generic Software Testing Process as the Target of Inspection

- Investigated state of practice
- Performed unstructured interviews with employees of an Austrian-based systems integrator to design a generic testing process
- Reviewed the process together with a software quality consultancy
- Aligned the process to the ISO/IEC/IEEE 29119 [2] series of standards

Overview

- Framework supports the VDI/VDE 2182 [7] guideline
- Ontological modeling approach
- Flexible assessment (scope)
- Combination of STRIDE [6] and attack–defense trees (ADTrees) [4]
- Automated generation of ADTrees
- Open-source prototype: `https://github.com/sbaresearch/adtgenerator`

Security Modeling Approach

- STRIDE: 6 categories of security threats used to build threat trees [6] that are included in the knowledge base
- ADTrees [4]: Attack trees extended by defense measures
- Description and formalization of various threat scenarios
- Automated generation of ADTrees, which can be imported into ADTool [3]
- Development of SPARQL queries to extract valuable security information from knowledge base (e.g., STRIDE threats to assets)

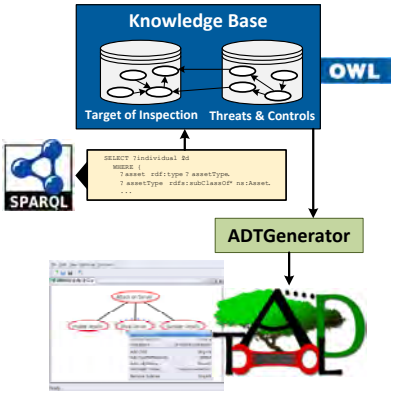


Figure 2: High-level overview of analyzing security risks in a semi-automated manner (ADTool illustrations taken from [3]).

Evaluation

- Two-step process: Tool selection step according to [5] and tool evaluation
- Considered 10 tools, two of which were extensively evaluated
- **Results:** Provides valuable support for security analyses, but needs to be improved to facilitate the structure analysis

Conclusion

- Designed a **generic software testing process for industrial automation applications** to define the target of inspection
- Proposed a framework that enables a **flexible, semi-automated security analyses**
- Adaptation to other engineering activities possible
- Developed a **prototype: ADTGenerator** (generation of ADTrees)
- SPARQL queries and ADTool [3] further support the analysis

Outlook

- **Automating risk identification** based on engineering data
- **Security modeling** extension for AutomationML (AMLsec)
- Detection of **vulnerabilities in plant structure** (e.g., attack graph generation), **consequences** of potential attacks, **business impact** analysis
- **Dynamic security risk analysis** methods for CPSs
- **Digital-twin-based attack simulation** for risk analysis

[1] M. Eckhart, K. Meixner, D. Winkler, and A. Ekelhart. Securing the testing process for industrial automation software. *Computers & Security*, 85:156 – 180, 2019.
[2] ISO/IEC/IEEE 29119-1. Software and systems engineering – software testing – part 1: Concepts and definitions, 2013.
[3] B. Kordy, P. Kordy, S. Mauw, and P. Schweitzer. ADTool: Security analysis with attack–defense trees. In K. Joshi, M. Siegle, M. Stoelinga, and P. R. D'Argenio, editors, *Quantitative Evaluation of Systems*, pages 173–176, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
[4] B. Kordy, S. Mauw, S. Radomirović, and P. Schweitzer. Foundations of attack–defense trees. In P. Degano, S. Etalle, and J. Guttman, editors, *Formal Aspects of Security and Trust*, pages 80–95, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
[5] R. M. Poston and M. P. Sexton. Evaluating and selecting testing tools. *IEEE Software*, 9(3):33–42, May 1992.
[6] A. Shostack. *Threat Modeling: Designing for Security*. Wiley Publishing, 1st edition, 2014.
[7] VDI/VDE 2182-1. Sheet 1: IT-security for industrial automation - general model, 2011.

This research was further funded by the FFG under the industrial PhD program (grant no. 874644). Moreover, the financial support by the Christian Doppler Research Association, the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

SmartIdentification
Secure Identification through Mobile Devices

Markus Maier

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

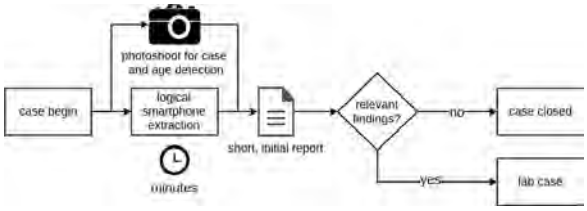
Problem & Motivation

- **Problems**
  - **Human trafficking** is a huge problem for **national security**.
  - Many people rescued **have no identificational documents**, but **most carry a smartphone**.
  - Sending smartphones to a lab for analysis **takes too much time** and **removes most forms of communication** for these people in distress.
- **Research Tasks**
  - **Identification of individuals**
    based on data found on a person's smartphone
  - **Detection of unattended minors**
    through analysis of images taken in the field and other sources
  - **Analysis of trends and used routes**
    by anonymizing and aggregating available location data
  - **Gather information about the trafficker**

Testing and Integration

The tests are heavily integrated into the ongoing efforts of both the **BM.I (Austrian Ministry of the Interior)** and the **Bundespolizei (German federal police)** against human trafficking. This allows to **rapidly integrate and fix features**, so that officers working with the software get the **optimal solution**. Further, this makes it possible to use **real data for tests** and correctly adjust the used technologies for detection without having access to the data.

In the Field



Officers in the field need a **quick overview** of the data on the phone. Primary focus is on **identifying documents** like passports; the extraction also contains **location data**, which is important for strategic evaluation like **trend and route analysis**.

In the Lab

In the lab, an **automated and fast routine** is beneficial to keep devices as briefly as possible. Further, current processes are mostly manual, thus an **extensive preliminary report** can **free resources** for more and faster analyses.

Keypoints

- **Various data sources from the provided smartphone**
  Phone numbers, contacts, text messages, images, device-specific data, location data, documents, connected WiFi
- **Techniques to spot relevant data while omitting irrelevant**
  Machine Learning as in image recognition, text and dialect detection, document detection (e.g. passports)
- **Age estimation based on images**
  Experimental checks to determine the age range of an individual (below 13, below 18, above 18).
- **Cross-checks between multiple cases**
  Finding common identifieres to highlight potential traffickers, e.g., one and the same phone-number on n analysed smartphones.
- **Testing during the development**
  Both the **BM.I** and the **Bundespolizei** heavily test the application in the field to ensure features work as intended and help effectively.
- **Extensive ethical and legal guidance**
  Established guidelines how this project can effectivly protect the indiviuals' privacy while providing valuable insights.

Trends and Routes

Through **manual export**, strategic evaluation can be provided based on **location data from multiple cases**. This data is anonymized through various techniques.

Figure 1: Demonstration of the trend and route analysis based on generated data

Conclusion

- SmartIdentification reduces the time needed per case by **providing information**, but **leaving the decision responsibility** with the officer in charge.
- People in distress get their smartphones back in a **more reasonable time**, as an extensive analysis in the lab is less often necessary.
- Valuable location data can be collected, providing better understanding of **trends during transit and the used routes**.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

WIP: Flashes of Lightning
Recovering Payment Channel Data from the Blockchain

Andreas Kern

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Information gathering in P2P networks can be a challenging task if reliable data is required. It is non-trivial to verify completeness even with mitigation techniques like constant monitoring from multiple points.

Observing the Lightning Network from the Blockchain will provide future work with reliable data.

- Can data about the Lightning Network reliably restored from onchain data?
- What kind of relevant data is stored onchain?
- How reliable is the recovery process?
- Which data cannot be reconstructed?
- Does the recovered data support measurements of the live Lightning Network data?

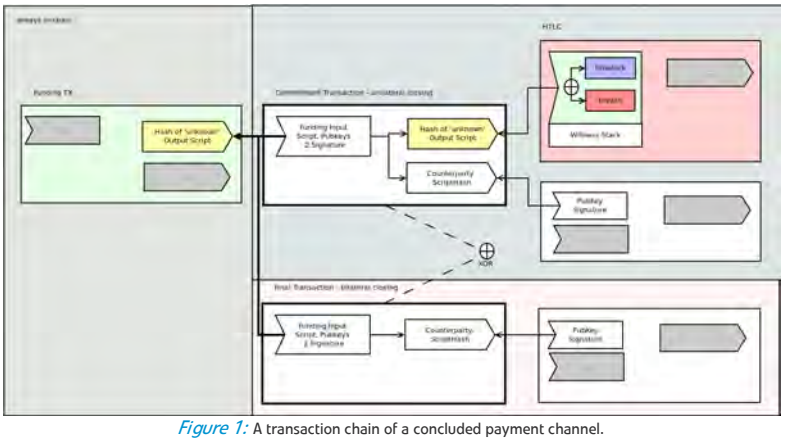


Figure 1: A transaction chain of a concluded payment channel.

Methodology

The data extraction happens as follows:

1. Data Acquisition with *bitcoind* in **JSON** format
2. Parsing the segwit stack with an **EBNF**-based parser written in **Python** to extract the relevant parts from the script and other items in the block
3. Transform data into **CSV** format
4. Filter and Visualize data with **R**

Challenges

- Pay2ScriptHash constructions reveal the actual transaction only after being spent.
- Transaction formats change and therefore onchain data can look very different from the examples given in the specification (BOLT vs Bech32).
- EBNF is an inadequate tool to parse length-prefixed data.
- JSON files are very slow compared to CSV data.
- Existing fields on the Bitcoin blockchain were repurposed to enable new functionalities without being renamed. Rules on how to interpret data are sometimes unexpected and hard to grasp, e.g. *nLocktime* changed its function and furthermore can represent a block height or a timestamp depending on the size of the value.

Extracted Data

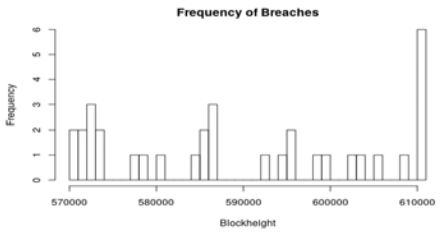


Figure 2: Frequency of Breaches

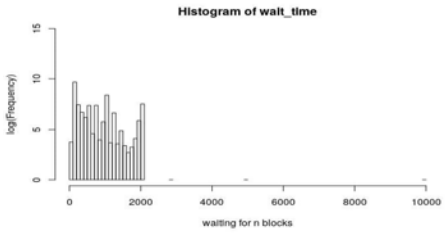


Figure 3: Channel timeout times in case of unilateral closing

Preliminary Results

- Valuable data can be extracted.
- Only completed channels that suffered a timeout or breach of contract can be detected safely.
- Length of time locks differ from default setting.
- Contract breaches do happen and Lighthouse services will be necessary for further adoption.

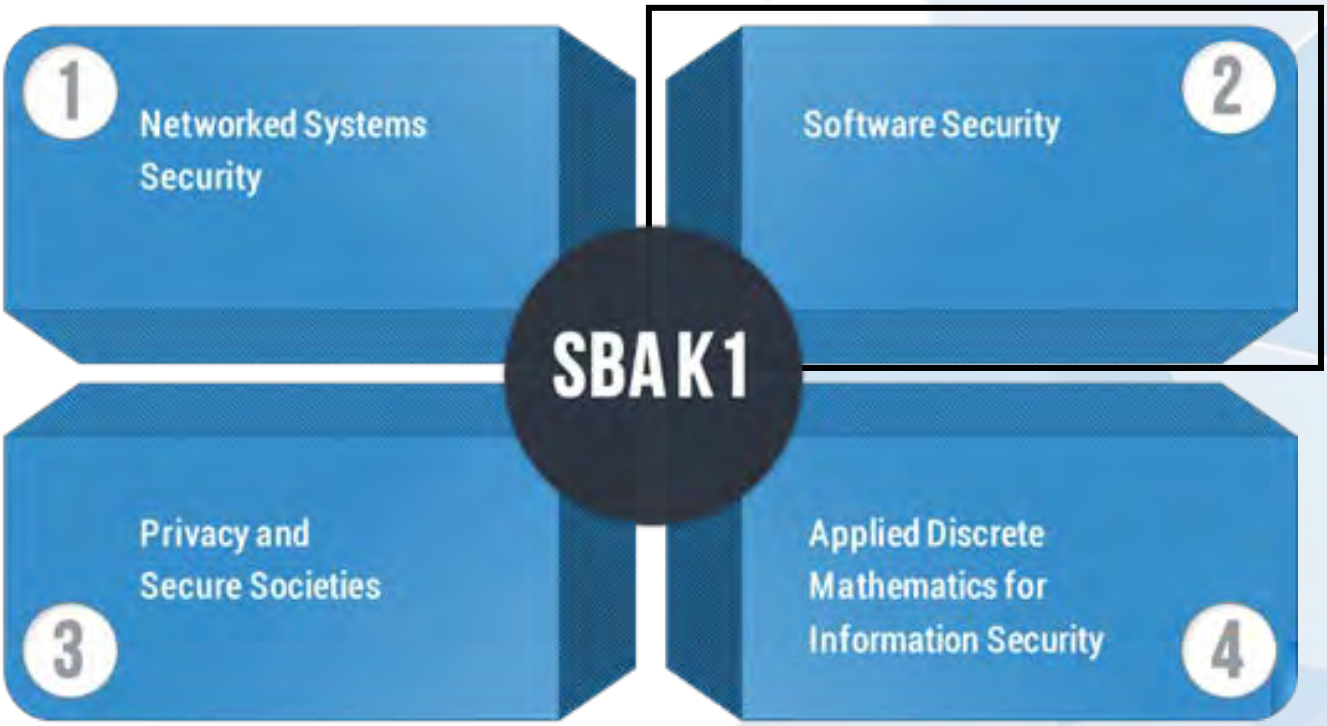
Future work

- How can this technique be applied to other blockchains?
- Are the implemented patterns universal?
- How to adapt the method to newly proposed channel types (Decker-Wattenhofer duplex, Decker-Russel-Osuntokun eltoo Channels, etc.)?
- What implications does the data (adoption rate, channel count, channel reliability, volume, etc.) have on privacy?

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
Die Stadt der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Software Security

AREA 2



Analytic Framework for Web-Application Penetration Testing

Stefan Haider

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Motivation & Contributions

Software security is a major concern in developing (web-) applications. Penetration testing is a best practice to find vulnerabilities before they can be exploited. Currently there are no tools or methods available that provide companies with measures on the quality of pentests. The quality and results of tests conducted by different security consultants can vary significantly, as they typically choose testing-time and testing-tools on their experience and budget. Metrics (based on empirical results) would help security consultants and companies to select and justify promising penetration testing settings.

Contributions

1. Design of a framework for collecting pentest information such as test results and relevant meta-data.
2. Implementation of a central, publicly available data repository for collecting pentest information.
3. Comparison of different clustering algorithms to find patterns in the collected data.
4. Evaluation of the prediction capabilities of the developed framework.

Data Collection

Table 1 shows a selection of the metrics and classifications used by the framework.

Metric/Classification	Category
Cyclomatic complexity	Application complexity
Lines of code	Application size
Experience	Team skill
Security awareness	Team skill
Code coverage	Test quality
Vulnerability severity	Vulnerability
Vulnerability density	Vulnerability
Methodology	Pentest quality
Pentester talent	Pentest quality
Application purpose	Application classification
Architectural style	Application classification
Technology stack	Application classification
Knowledge	Pentest classification
Location	Pentest classification
Tool-support	Pentest classification

Table 1: Selected metrics used by the framework

Data Clustering

Clustering is used to find patterns in the collected data, as our hypothesis states a correlation between the applications measures and the vulnerabilities identified during a pentest. To select a suitable clustering algorithm we:

1. use a randomized search strategy for parameter optimization for each algorithm,
2. compare the measures of the generated clusters,
3. and select the best suitable algorithm(s).

Data Generation

To evaluate the pattern detection ability of the framework, we use open-source projects hosted on GitHub for our evaluation.

1. Search for GitHub repositories with the keywords: "webapp", "cms", "ecommerce", "blogging".
2. Analyze these repositories with Sonarqube (https://www.sonarqube.org/) to generate measures regarding the sourcecode for the metrics listed in table 1.
3. Find publicly disclosed vulnerabilities in the National Vulnerability Database (NVD) and link them to the applications.
4. Use these vulnerabilities to estimate the result of a pentest.

Figure 1 shows an example of the generated measures during a sonarqube scan.

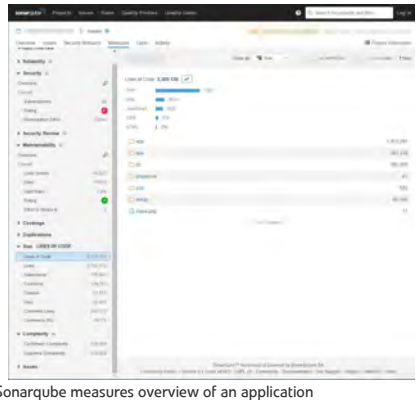


Figure 1: Sonarqube measures overview of an application

Expected Results

By finding clusters in the data collected via our defined framework we expect to optimize future pentest and predict the outcome of those. As input this method requires a static sourcecode analysis and simple metadata for the application such as (the technology stack, measured software team skills, etc.). Using this approach for the following optimizations is of particular interest.

- The minimum amount of time necessary to perform a pentest covering a given amount of vulnerabilites.
- The best tools used during a pentest for a certain application based on its metadata.

Additionally to the optimizations of the pentesting process the clusters can also be used for predicting vulnerabilities found during a pentest.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Contactless Side Channel Based Disassembling
A time shift resilient machine learning approach

Tobias Kovats, Georg Merzdovnik, Christian Kudera

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Side Channel Based Disassembling
Refers to the process of recovering the source code running on a device from one of its side channels, e.g., its power consumption or electromagnetic emanation.

This can be divided into two phases:

- **Profiling:** Behaviour of the chosen side channel is learned during execution of corresponding instructions.
- **Attacking:** Unknown instruction traces are reconstructed from side channel leakage.

Problems
Side channel obfuscation is predominantly caused by

- **Non-consecutive execution:** Instruction phases overlap due to processor pipeline.
- **Register value dependencies:** Side channel behaviour depends not only on instruction, but also on register values. [1]
- **Synchronization:** Attacker needs to be in-phase with the target device.
- **Background noise:** Peripherals and other unwanted radiation sources cause interference.

Our Approach

Contribution: We incorporate previously proposed approaches from secret key recovery [2] and image recognition [3] in building an electromagnetic side channel based disassembler with enhanced time shift resilience to eliminate the need for synchronization.

- **Profiling:** To ensure independence of all other factors except the instruction itself, all registers are initialized with random values, and two random instructions are inserted before and after the target instruction, each working with random registers.
- **Data Augmentation:** Several data augmentation methods are evaluated. The following aims to enhance time-shift resilience: Several trace recordings are concatenated; a window of length Γ and a uniform random offset $\Gamma' \epsilon [-\sigma, +\sigma]$ from t_0 cut out.
- **Attacking:** For performance evaluation, random instruction traces are generated and reconstructed.
- **Sliding Window:** A window of length Γ slides over the full trace using a predefined overlap. The resulting windows are fed into the classifier.
- **Model:** A one-dimensional convolutional neural network similar to the VGG16 implementation is used for classification.

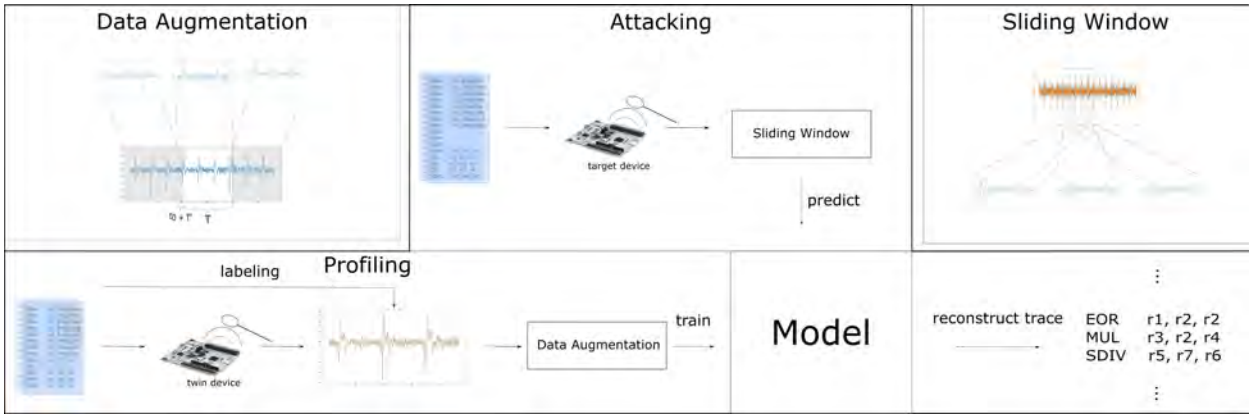


Figure 1: Algorithm schematics and close up visualizations of the preprocessing algorithms

Application

- **Reverse Engineering:** IP theft, localization of critical code segments, code recognition and code flow analysis
- **Monitoring:** Detection of deviations from regular execution, realtime code execution tracking
- **Complex Embedded Systems:** Non-intrusive application to systems that don't allow time synchronization, e.g., programmable logic controllers

Preliminary Conclusions*

- **Complex model architectures** and high computational resources are necessary to cope with side channel obfuscations.
- **Leakage cartography** of local electromagnetic emanations shows high gradient.
- **Promising results** using a reduced instructions set have already been obtained.

*This is a work in progress and final conclusions are to follow.

[1] Thomas Eisenbarth, Christof Paar, and Björn Weghenkel. Building a side channel based disassembler. *Transactions on Computational Science*, 10:78–99, 01 2010.
[2] Eleonora Cagli, Cecile Dumas, and Emmanuel Prouff. Convolutional neural networks with data augmentation against jitter-based countermeasures – profiling attacks without pre-processing –. Cryptology ePrint Archive, Report 2017/740, 2017. https://eprint.iacr.org/2017/740.
[3] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640, 2015.
[4] Valence Cristiani, Maxime Lecomte, and Thomas Hiscock. A bit-level approach to side channel based disassembling. In Sonia Belaid and Tim Güneysu, editors, *Smart Card Research and Advanced Applications*, pages 143–158, Cham, 2020. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
Das Kompetenzzentrum SBA Research wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMVIT, BMDW, das Land Wien gefördert. Das Programm COMET wird durch die FFG abgewickelt.

Control-Flow Integrity: Precision, Security, and Performance

Nathan Burow, Scott A. Carr, Joseph Nash, Per Larsen, Michael Franz, Stefan Brunthaler, Mathias Payer

Motivation & Contributions

Memory corruption errors in C/C++ programs remain the most common source of security vulnerabilities in today's systems. Control-flow hijacking attacks exploit memory corruption vulnerabilities to divert program execution away from the intended control flow. Researchers have spent more than a decade studying and refining defenses based on Control-Flow Integrity (CFI); this technique is now integrated into several production compilers. However, so far, no study has systematically compared the various proposed CFI mechanisms nor is there any protocol on how to compare such mechanisms.

Contributions

1. A systematization of CFI mechanisms with a focus on discussing the major different CFI mechanisms and their respective trade-offs.
2. A taxonomy for classifying the underlying analysis of a CFI mechanism.
3. Presentation of both a qualitative and quantitative security metric and the evaluation of existing CFI mechanisms along these metrics.
4. A detailed performance study of existing CFI mechanisms.

Proposed Classifications

Control-Flow Transfers

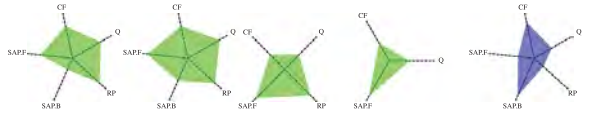
- CF.1: backward control flow
- CF.2: forward control flow using direct jumps
- CF.3: forward control flow using direct calls
- CF.4: forward control flow using indirect jumps
- ...

Static Analysis Precision

- SAP.F.0: No forward branch validation
- SAP.F.1a: ad-hoc algorithms and heuristics
- SAP.F.1b: context- and flow-insensitive analysis
- SAP.F.1c: labeling equivalence classes
- ...

Qualitative Security Guarantees

We conducted a qualitative analysis of prior work and proposed CFI implementations. Figure 1 presents the results for selected open-source LLVM-based implementations.



(a) MCFI [Niu and Tan 2014a] (b) πCFI [Niu and Tan 2015b] (c) IFCC [Tice et al. 2014] (d) LLVM-CFI-3.7 (2015) (e) Lockdown [Payer et al. 2015]

Figure 1: Comparison: control flows (CF), quantitative security (Q), reported performance (RP), static analysis precision: forward (SAP.F) and backward (SAP.B). Color coding of CFI implementations: binary are blue, source-based are green.

Performance

Our performance experiments show that recent compiler-based CFI mechanisms have mean overheads in the low single-digit range.

Benchmark	LLVM-CFI (3.9)	CFGuard	πCFI	πCFI (ntc)
400.perlbench(C)	2.4		8.2	5.3
401.bzip2(C)	-0.7	-0.3	1.2	0.8
403.gcc(C)	CF		6.1	10.5
429.mcf(C)	3.6	0.5	4.0	1.8

Table 1: Measured CFI Performance Overhead (%) on the SPEC CPU2006 Benchmarks (excerpt)

Quantitative Security Guarantees

One possible metric for assessing how much security a CFI mechanism provides is the product of the number of equivalence classes (ECs) and the inverse of the size of the largest class (LC).
We conducted three different quantitative evaluations in line with our proposed metric for evaluating the overall security of a CFI mechanism. Figure 2 shows the number of ECs for the five CFI implementations that we evaluated as well as their subconfigurations.

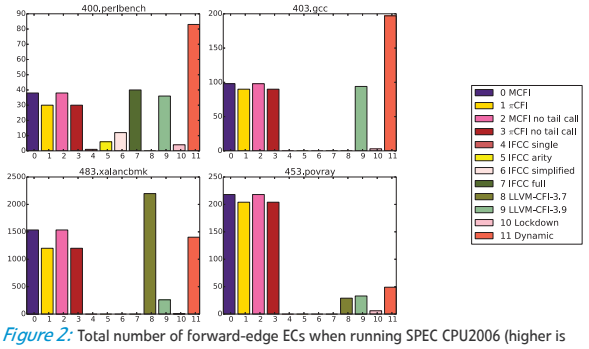


Figure 2: Total number of forward-edge ECs when running SPEC CPU2006 (higher is better).

Open Problems

- Most existing CFI implementations use ad hoc, imprecise analysis techniques. Future work in CFI should use flow-sensitive and context-sensitive analysis for forward edges. On backward edges, we recommend shadow stacks.
- Quantifying the incremental security provided by CFI or any other security mechanism is an open problem. However, a large adversarial analysis study would provide additional insight into the security provided by CFI.

Research Frontiers

- Protecting Operating System Kernels
- Protecting Just-in-time Compiled Code
- Protecting Interpreters
- Protecting Method Dispatch in Object-Oriented Languages

Nathan Burow, Scott A. Carr, Joseph Nash, Per Larsen, Michael Franz, Stefan Brunthaler, and Mathias Payer. Control-flow integrity: Precision, security, and performance. *ACM Comput. Surv.*, 50(1), April 2017.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Data Poisoning Attacks in Federated Learning

Rudolf Mayer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

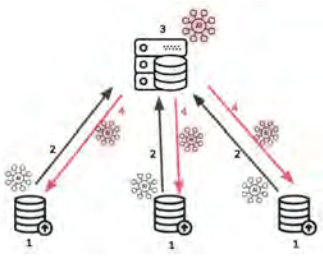
Federated Learning promises advances over centralized learning:

- No need to exchange distributed data for collaborative learning
  - Alleviating many risks and obstacles related to data privacy
- Computing resources at the data holders can be utilized, thus distributing the computation

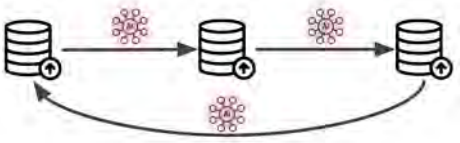
Federated Learning is however still (or even more) exposed to adversarial attacks. We evaluate to what extend an attacker can disturb the training process to successfully embed a backdoor in the common model.

Federated Learning

Rather than moving the data to the model, Federated Learning is based on the principle of creating a model where the data is generated. Two different architectures of federated machine learning can be distinguished.
Parallel Federated Learning: Each training round consists of several steps: The clients train models based on their local data (1), which are send to the aggregation server (2). There, the local models are combined (e.g. by averaging the models' parameters) (3). Finally, they are distributed back to the clients (4).



Sequential Federated Learning (aka cyclic incremental learning): A client trains its model locally, and sends it to the next client for further training. This does not require a central aggregation process.



Poisoning Attacks

Despite obvious benefits, the distributed nature of Federated Learning enables new attack vectors for adversaries. Backdoor attacks are an attack targeting the model's integrity during the training phase. According to this strategy, an adversary poisons the training data by adding samples containing a certain pattern (the so-called "backdoor"). The goal is to trigger malicious behavior on data containing this pattern during the deployment phase. *Note*: The appearance of the backdoor patterns as such is not a primary concern. While the created backdoors are noticeable, they are chosen to be unsuspicious as they naturally occur in the selected data.

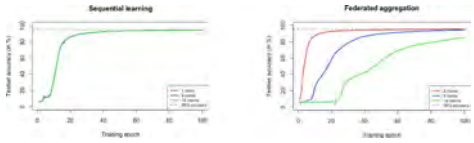
(a) Original image (Yale Face dataset) *(b)* Backdoor pattern „glasses" *(c)* Original image (Traffic Sign dataset) *(d)* Backdoor pattern „black square"

Goals

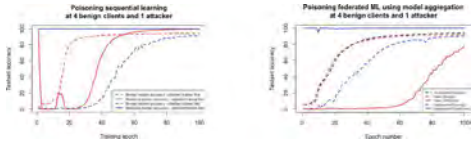
We compare federated machine learning to central machine learning on aspects such as the number of participants in the network, or the ability of handling non-independent and non-identically distributed (non-i.i.d.) data. We measure the effectiveness of different ML models with common metrics such as test set accuracy. Furthermore, we perform backdoor attacks in Federated Learning settings to gain insight into the impact on effectiveness by varying properties such as the pattern's appearance (size, shape or color), attack strategies, or varying numbers of attackers.

Results

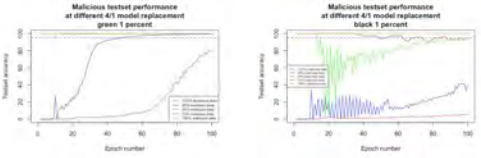
Tested on i.i.d. data, the number of clients has no influence on the effectiveness of the machine learning model in the case of sequential learning. However, in the case of federated (parallel) aggregation, a higher number of clients leads to a slower convergence of the global model. Sequential learning on non-i.i.d. data suffers from catastrophic forgetting, meaning that data trained in early stages is underrepresented in the resulting model. Federated averaging also suffers from reduced performance on sparsely known classes.



Backdoor attacks can be successfully introduced in both federated settings, shown below in networks consisting of 4 benign clients and 1 attacker. In sequential learning, the point of time the attacker participates in the learning cycle has a big impact on the performance. In a federated aggregation setting, especially the model replacement strategy [1] leads to a high effectiveness on benign and malicious test data.



We can further see that black color is less effective for the backdoor: the only backdoor attack considered as successful is when malicious client uses 50% poisoned data. If we use only 25%, the performance of the malicious test set is low, and the attack is not successful



Conclusions

We evaluated different types of Federated Machine Learning techniques regarding effectiveness on benchmark datasets for image classification. Federated ML offers advantages regarding privacy and utilisation of resources, but opens up new attack vectors for adversaries. We designed and implemented strategies for backdoor attacks and were able to confirm that Federated ML is highly susceptible to these attacks. Future work will put an emphasis especially on defence strategies.

E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov. How To Backdoor Federated Learning. In *23rd International Conference on Artificial Intelligence and Statistics (AISTATS)*, Palermo, Italy, June 2020.
F. Nuding and R. Mayer. Poisoning Attacks in Federated Learning – an Evaluation on Traffic Sign Classification. In *10th ACM Conference on Data and Application Security and Privacy (CODASPY)*, New Orleans, U.S., March 2020.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Federated Machine Learning in Privacy-Sensitive Settings

Anastasia Pustozerova

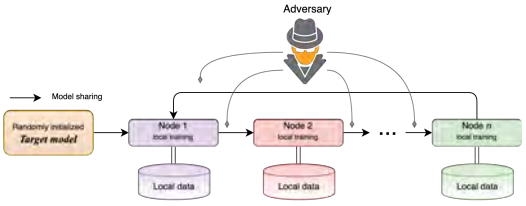
COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Federated learning allows performing machine learning over distributed data while *preserving privacy* of data owners. Each data holder independently and locally trains a machine learning model on her own data and then shares the model with other participants of the federated learning process, so other parties can proceed training on their own data, or aggregate several models to a global one. Federated learning addresses the issue of *data locality and sensitivity* and also enables using *computational power of distributed systems*, closer to the place where the data is originating. However, models, which are exchanged during the federated learning process, can leak information about their training data. In this work, we

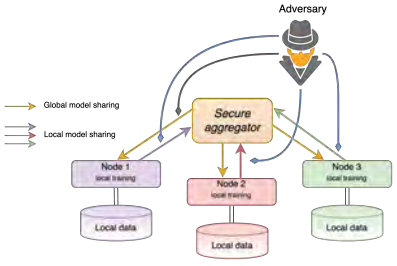
- evaluate privacy risks in federated learning by performing membership inference attack,
- propose mitigation strategies to improve privacy properties of federated learning,
- develop guidelines for federated learning allowing to maintain effectiveness of the models while preserving privacy of the data.

Parallel federated learning vs. Sequential Federated learning



Sequential federated learning

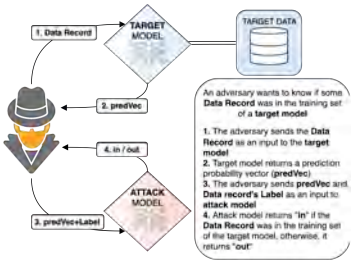
A randomly initialized model is locally trained at the first client and then passed to the next node in the sequence. After completing a full round of ***n*** nodes, the model is passed again to the first node for repeating the training process.



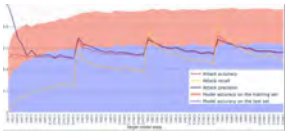
Parallel federated learning

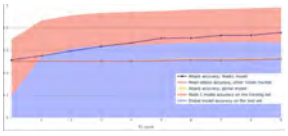
An *aggregator* initializes a global model with random weights and shares it to every node in the setting. Each node trains the model in parallel on its local data and then returns it to the secure aggregator. From the locally trained models, a new global model is aggregated and shared to the clients for the following training cycle.

Membership inference attack



Attack evaluation

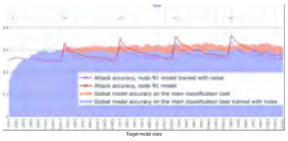


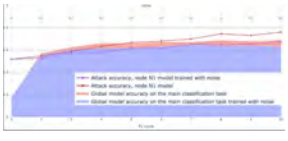


In **sequential federated learning**, the attack on node N1 data has the highest accuracy while performing membership inference on the shared model right after training at node N1.

In **parallel federated learning**, the attack on node N1 data has higher accuracy while performing membership inference on the model trained locally at N1, than attacking global model, or local models from other nodes.

Mitigation evaluation





In both **sequential** and **parallel** federated learning adding, noise to the training data allows to mitigate the risks of membership inference attack. However, the noise should be properly chosen to not cause loss in effectiveness of the global model on the classification task.

Conclusion

- Federated learning allows to avoid data transferring while training machine learning models without loosing in models effectiveness
- The models shared during federated learning process can leak information about their training data, e.g. when attacker performs membership inference attack
- The membership inference attack accuracy can be reduced by adding noise to the training data.

References:

1. A. Pustozerova, and R. Mayer, "Information Leaks in Federated Learning", Workshop on Decentralized IoT Systems and Security (DISS), San Diego, CA, USA, 2020
2. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data.," 20th International Conference on Artificial Intelligence and Statistics, (AISTATS), Fort Lauderdale, FL, USA, 2017.
3. S. Truex, L. Liu, M. Gursoy, L. Yu, and W. Wei, "Demystifying membership inference attacks in machine learningas a service,"IEEE Transactions on Services Computing, 2019.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
Das Kompetenzzentrum SBA Research wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMVIT, BMDW, das Land Wien gefördert. Das Programm COMET wird durch die FFG abgewickelt.

From Hack to Elaborate Technique
A Survey on Binary Rewriting

Matthias Wenzl, Georg Merzdovnik, Johanna Ullrich, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Survey Motivation

Binary Rewriting?

- Software is often distributed in binary form or needs to be changed during runtime.
- Originally inspired by the need to change parts of a program while software is executed.
- Nowadays, evolved into a plethora of approaches with different application domains (e.g. Emulation, Observation, Optimization, Hardening).

Problem

- A plethora of different approaches and methods has led to the development of many different tools.
- However, because of this, it is not always easy to identify the right tool for the problem at hand.
- Additionally, the availability of tools and methods for specific purposes is not well studied.

Rewriting at a Glance

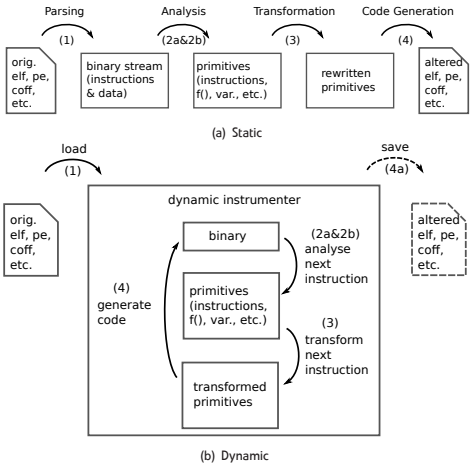


Figure 1: Required steps to apply binary rewriting in principle.

4 Steps of Rewriting

1. **Parsing:** Extract instruction and data stream from binary objects for further analysis
2. **Analysis:** Provides information on buidling blocks (e.g., disassembly, structural recovery or label, symbol and data type extraction)
3. **Transformation:** Prepare instrumentation points and define alterations (e.g., to instructions or control flow)
4. **Code Generation:** Apply the intended changes into the binary of interest in a way to keep it executable

Transformations

Static perform alterations directly at instrumentation point (e.g. during link time)

Dynamic Able to perform changes at instruction granularity during runtime

Minimal-invasive operations on branch granularity, by redirecting control flow to newly generated code

Full-translation transform binaries at any instruction, but require lifting into Intermediate Representation (IR)

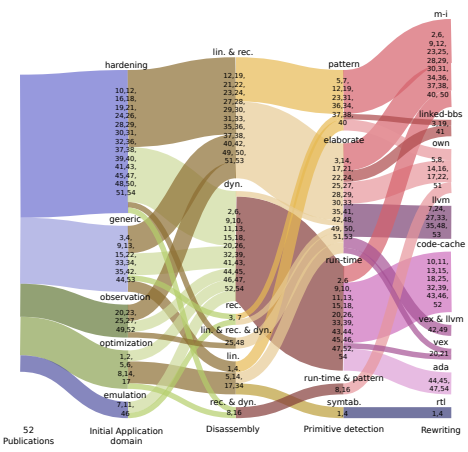


Figure 2: Sankey diagram further categorizing the publications listed in [1]. The adjacent row depicts the tool's application domain, followed by its used disassembly, structural recovery, and transformation strategy.

Conclusion

- Full-translation-based schemes allow for application of reasoning approaches due to the more abstract representation of the binary under investigation.
  - Currently only semantic equivalent lifters are available, which are sufficient for many applications.
  - Scenarios like altering timing sensitive applications, performance optimization for throughput-oriented programs, or rewriting software with real-time requirements would greatly benefit from instruction equivalent lifters.
- The x86 architecture is still the primary target for binary rewriting applications, but other architectures like ARM and MIPS draw more and more interest.

[1] Matthias Wenzl, Georg Merzdovnik, Johanna Ullrich, and Edgar Weippl. From hack to elaborate technique—a survey on binary rewriting. *ACM Computing Surveys (CSUR)*, 52(3):1–37, 2019.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Mitigating Rowhammer Attacks with Software Diversity

Manuel Wiesinger

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

The Rowhammer vulnerability

- Hardware vulnerability
  - Allows to change the state of bits in primary memory
  - Cannot be fixed by applying simple software updates
- Affects almost any modern computer system
  - Mobile phones
    - Malicious apps can "root" devices
  - Cloud infrastructure
    - Attackers can gain control of other virtual machines, running on the same physical host
  - Consumer and office computers
    - Allows to gain superuser privileges (even from web browsers) [3]

Attackers can exploit Rowhammer by tricking operating systems to store critical data at memory locations vulnerable to bit flips. These tricks work, because central memory management algorithms work predictably.

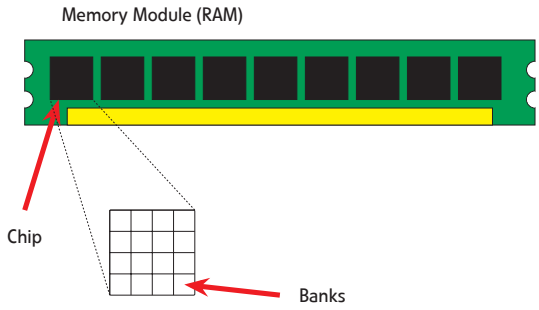


Figure 1: The Rowhammer vulnerability allows to change bits in memory banks.

State of the Art

Defenses suggested by previous work have several problems

- Require hardware exchange [4]
- Aren't effective against all Rowhammer variants [2]
- Dramatically increase boot times and physically damage memory modules over time [1]

Methodology

- Analysis of existing (publicly known) attacks
- Evaluation of proposed defenses
- Implementation of a prototype for our novel defense against Rowhammer based attacks in the Linux kernel
- Evaluation of the prototype

Mitigating Rowhammer

Page Sacrifice

- Idea: Sacrifice some memory blocks for better security
  - Add a random padding before and after each memory block
  - Associate paddings with memory blocks that are actually used to store data
  - Free padding blocks when memory blocks are freed
- Implementation
  - Implement a proof of concept implementation for the Linux kernel
  - Provide run time configurable parameters (via the `sysctl` interface)

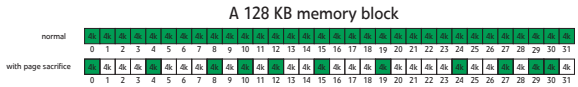
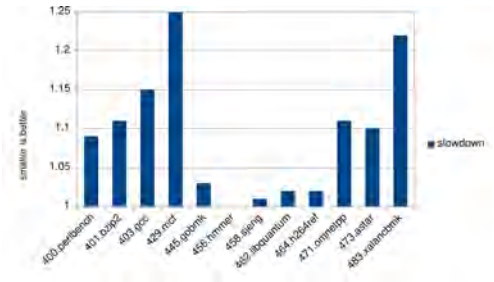


Figure 2: Example of memory block with smaller sacrificed memory blocks using rather extreme parameters

SPEC CPU 2006 Evaluation



Results

- Attackers cannot predict the padding and fail to store memory posterblocks containing critical information at vulnerable locations
- As shown in the figure above the execution slowdown of page sacrifice is negligible for many use cases.
- Experiments indicated that our defense requires 2-3 times more memory using reasonable parameters.

Franz Ferdinand Brasser, Lucas Davi, David Gens, Christopher Liebchen, and Ahmad-Reza Sadeghi. Can't touch this: Practical and generic software-only defenses against rowhammer attacks. *CoRR*, abs/1611.08396, 2016.
Daniel Gruss, Moritz Lipp, Michael Schwarz, Daniel Genkin, Jonas Juffinger, Sioli O'Connell, Wolfgang Schoechl, and Yuval Yarom. Another flip in the wall of rowhammer defenses. In *2018 IEEE Symposium on Security and Privacy, SP 2018, Proceedings, 21-23 May 2018, San Francisco, California, USA*, pages 245–261, 2018.
Daniel Gruss, Clémentine Maurice, and Stefan Mangard. Rowhammer.js: A remote software-induced fault attack in javascript. DIMVA 2016, 2016.
M.Ghasempour, M. Lujan, and J.Garside. Armor: A run-time memory hot-row detector, 2015.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
Ein Fonds der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
Das Kompetenzzentrum SBA Research wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMVIT, BMDW, das Land Wien gefördert. Das Programm COMET wird durch die FFG abgewickelt.

Sendo: End-to-end Encrypted and Easy-to-use File Transfer

Thomas Konrad, Michael Koppmann and Andreas Boll

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

When sharing sensitive data with people outside the own ogranization, it is hard to make a good technology choice for multiple reasons.

- Email encryption solutions such as PGP or S/MIME are not widely available,
- they lack support in some widely-used email clients, and
- they are generally hard to use for non-IT people.

The idea was to create an easy-to-use web application where files are encrypted in the browser, uploaded to a server and securely shared via a link. Files are automatically deleted after a configurable time period to minimize the attack surface.

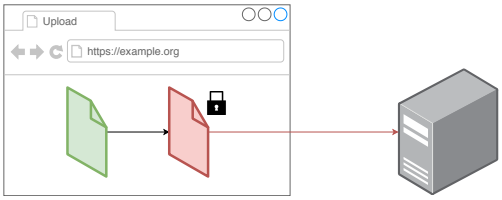


Figure 1: Files are encrpyted before they are uploaded to the server.

The Design

The two essential elements of the overall system design are **encryption** and **link sharing**.

Encryption

The files are encrypted as follows:

- A master key is generated in the browser using the Web Crypto API.
- The file is split up into small chunks using the FileReader API.
- Each chunk is encrypted with AES-GCM and a unique IV.

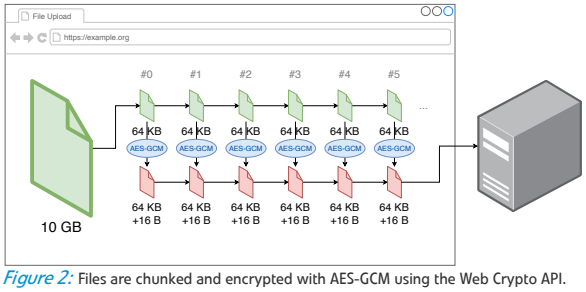


Figure 2: Files are chunked and encrypted with AES-GCM using the Web Crypto API.

Link Sharing

To protect the link, the master key is encrypted with a password which is also generated by the secure random number generator of the Web Crypto API. The master key is then encrypted with this password. PBKDF2 is used to significantly slow down brute-force attacks on the password.

In order to avoid that the encrypted master key is ever sent to the server, it is stored in the fragment part of the URL. The fragment only stays in the browser and is never transferred via a HTTP request.

The system enforces that the link and the password are transferred via two separate channels.

Modern Browser APIs

Sendo uses many cutting-edge browser APIs for its functionality. It generates random numbers and encrypts files using the **Web Crypto API**, splits files with the **FileReader API**, and uses **Service Workers** and the **ReadableStream API** to stream file data.

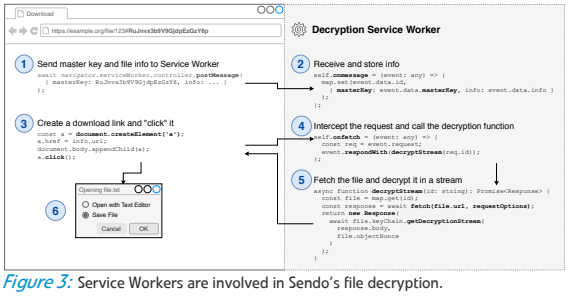


Figure 3: Service Workers are involved in Sendo's file decryption.

Figure 4: Sendo is implemented as a web browser application.

Conclusion

Sendo aims to provide an **easy-to-use** way of **securely sharing files** with people outside the own organization. It uses **cutting-edge browser APIs** for encryption and key management and automatically deletes files after a configurable time period. It enforces that the link and the password are sent via **different channels** so that the likelihood of a successful attack is very low.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
Ein Fonds der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
Das Kompetenzzentrum SBA Research wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMVIT, BMDW, das Land Wien gefördert. Das Programm COMET wird durch die FFG abgewickelt.

Speculator: A Tool to Analyze Speculative Execution Attacks and Mitigations

Andrea Mambretti, Matthias Neugschwandtner, Alessandro Sorniotti, Engin Kirda, William Robertson, Anil Kurmus

Problem & Motivation

Speculative execution attacks (SEAs) exploit vulnerabilities at a CPU's microarchitectural level, which, until recently, remained hidden below the instruction set architecture. The full breadth of this class of attacks is not entirely understood, owing to the fact that details about the microarchitectural level of modern commercial CPUs are not publicly available. Hence, the research community cannot provide complete answers to questions about the existence of new attacks and the effectiveness of defenses.

Requirements

- A tool that inspects microarchitectural state directly to analyze speculative execution attacks.
- Such a tool should be beneficial to gain confidence in speculative execution mitigations being properly implemented and enabled.

Speculator

Speculator is a method and tool to reverse-engineer the behavior of CPUs (across architectures and vendors) in order to build a deeper understanding of speculative execution. Speculator aggregates the relevant sources of information, including CPU performance counters and model-specific registers.

Objectives

- Accurately measure microarchitectural state attributes associated to the speculative portion of the execution of user-supplied snippets of code.
- Observe (i) which parts of the snippet are speculatively executed, (ii) what causes speculative execution to start and stop, (iii) what parameters affect the amount of speculative execution, (iv) how do specific instructions affect the behavior of speculative execution, (v) which security boundaries are effective in the prevention of speculative execution, and (vi) how consistently CPUs behave.
- Support multiple platforms for inference of generalizable facts about speculative execution.

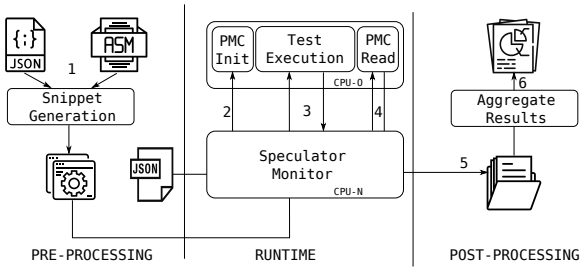


Figure 1: Architecture of Speculator

Microarchitectural behavior of modern CPUs

Using Speculator, we explored the microarchitectural behavior of a selection of modern CPUs (Intel Haskwell, Broadwell, Skylake, Kaby Lake, Coffee Lake, and AMD Zen) to investigate several speculative execution properties, e.g.,:

- Return Stack Buffer Size (successfully verified)

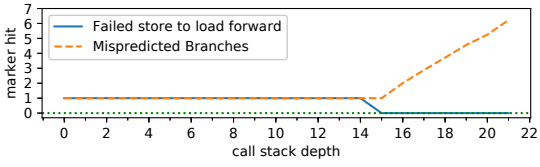


Figure 2: Return Stack Buffer test on Kabylake

- Nesting Speculative Execution (works)
- Speculative execution across system calls (does not span across system calls)
- Flushing the Cache (`clflush` has no effect during speculation)
- Speculation window size (measured for indirect branches, indirect control flow transfers and store to load forwards)

Table 1: Speculation window of an indirect control flow transfer (in cycles)

Indirect branch target location	Broadwell	Skylake	Zen
Register	28	22	24
Cached memory	41	34	35
Uncached memory	154	303	301

- Stopping Speculative Execution (verified `lfence` as barrier for speculative execution)
- Executable Page Permission (execute permissions are enforced)

Attacks & Mitigations

We investigated new techniques to exploit speculative execution attacks and mitigations. With snippets of code we verified the behavior during speculation and if an attack might be feasible through those instructions.

- Split Spectre: aims to lower the requirements for a Spectre v1 attack with the attacker being able to provide the second of the two array accesses required. Could not be exploited in SpiderMonkey 52.7.4.
- Branch Target Injection (BTI): analyzed the feasibility of BTI poisoning between co-located processes. Success rate up to 82% over a thousand runs on Skylake and Kaby Lake, up to 55% on Coffe Lake and Broadwell, and no success on AMD Ryzen.
- Mitigations: tested Spectre v2 mitigations available in the kernel (STIBP, IBRS and IBPB). While the tested countermeasures are effective, default settings do not enable them.

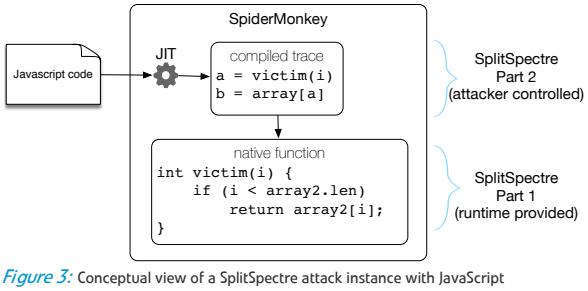


Figure 3: Conceptual view of a SplitSpectre attack instance with JavaScript

Andrea Mambretti, Matthias Neugschwandtner, Alessandro Sorniotti, Engin Kirda, William Robertson, and Anil Kurmus. Speculator: A tool to analyze speculative execution attacks and mitigations. In *Proceedings of the 35th Annual Computer Security Applications Conference*, ACSAC '19, page 747–761, 2019.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Synthetic Data
Utility Evaluation for Machine Learning

Markus Hittmeir, Andreas Ekelhart, Rudolf Mayer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Due to technological advances, an increasing amount of micro-data, i.e., data that contains information about individuals, is collected. To comply with ethical and legal standards, data holders have to take privacy-preserving measures. Traditional concepts like k-Anonymity and Differential Privacy prevent intruders from learning sensitive information about individuals in the dataset. An alternative is the generation of *synthetic data*, which usually consists of the following steps:

1. **Data Description**: The original data is used to build a model which comprises information about the distribution of attributes and correlations between them.
2. **Data Generation**: The model is used to generate data samples. The global properties of the resulting synthetic dataset are similar to the original, but the samples do not represent real individuals.

Ultimate Goal: Machine Learning models trained on synthetic data instead of the real data perform nearly as well. Simultaneously, the use of synthetic data reduces the risk of disclosure of sensitive information.

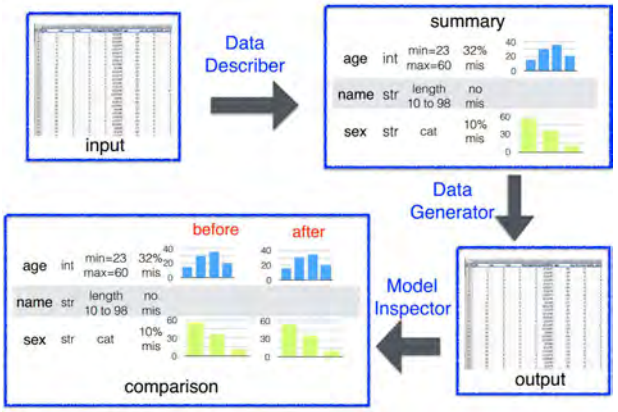


Figure 1: Workflow of the DataSynthesizer (Figure from Ref. 2)

Synthetic Data Generation Tools

- **Synthetic Data Vault (SDV)**[1]: This tool builds a model based on distribution estimates for each column and the covariance matrix for preserving correlations.
- **DataSynthesizer**[2]: The DataSynthesizer's model is based on a Bayesian network learned from the original data. For statistical disclosure control, the user is able to turn on Differential Privacy. In our experiments, we have evaluated the DataSynthesizer both with enabled (**DSP**) and disabled (**DS**) Differential Privacy.
- **synthpop (SP)**[3]: The default synthesis method is the CART algorithm. However, the user may specify a large number of parameters. Moreover, the implementation comes with its own function for statistical disclosure control.

Utility Evaluation for Classification

We performed classification tasks on standard benchmark data, such as the Adult Census dataset from the UCI Machine Learning Repository:

1. We split the data into a training (80 %) and a test (20 %) dataset.
2. We applied the synthesizers to the original training data to generate synthetic training data of equal length.
3. We trained machine learning classification models on both the original and the synthetic training data.
4. We compared the accuracy scores of these models on the test data.

General Findings

The **DS** and **SP** with standard settings achieve accuracy scores close to the models trained on original data. Using the **SDV** or the **DSP** usually leads to a loss of performance. Figure 2 shows an example, where **O** is the accuracy score of the original data and **B** the baseline score given by a Zero Rule classifier.

Figure 2: Accuracy scores of Logistic Regression on the Adult Census dataset

Utility Evaluation for Regression

Using a similar experimental setup as for classification and benchmark data such as the Bike Sharing dataset from Kaggle, we evaluated the utility of synthetic data for solving regression tasks. In such tasks, the target variable is continuous and the problem is not to predict a category, but a numerical value. The goal is to be as close to the sample's real value as possible.

General Findings

We compared the results of multiple utility measures, such as the mean average error (MAE) and the R2 score. In accordance with our analysis on classification tasks, the **DS** and **SP** with standard settings usually achieve scores close to the models trained on original data. Using the **SDV** or the **DSP** still leads to a performance loss. However, Figure 3 is an example of the **SDV**'s tendency to perform much better on regression than on classification problems, as its MAE is only slightly larger than the MAE for **SP** and **DS**.

Figure 3: MAE for Support Vector Regression on the Bike Sharing dataset

1. N. Patki, R. Wedge, K. Veeramachaneni, *The Synthetic Data Vault*, In: Proceedings of the 3rd DSAA, 2016.
2. H. Ping, J. Stoyanovich, B. Howe, *DataSynthesizer: Privacy- Preserving Synthetic Datasets*, In: Proceedings of the 29th SSDBM, 2017.
3. B. Nowok, G. M. Raab, C. Dibben, *synthpop: Bespoke Creation of Synthetic Data in R*, In: Journal of Statistical Software, 2016.
4. M. Hittmeir, A. Ekelhart, R. Mayer, *On the Utility of Synthetic Data: An Empirical Evaluation on Machine Learning Tasks*, In: Proceedings of the 14th ARES Conference, 2019.
5. M. Hittmeir, A. Ekelhart, R. Mayer, *Utility and privacy assessments of synthetic data for regression tasks*, In: Proceedings of the 7th IEEE Big Data Conference, 2020.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

WPSE: Fortifying Web Protocols via Browser-Side Security Monitoring

Stefano Calzavara, Riccardo Focardi, Matteo Maffei, Clara Schneidewind, Marco Squarcina, Mauro Tempesta

Motivation & Approach

Designing and implementing web protocols, such as OAuth 2.0, OpenID Connect, and SAML 2.0, is an error-prone task even for security experts, as witnessed by the large number of reported vulnerabilities. The main reason for this is that web protocols involve communication with a web browser, which is not aware of the existence of web protocols and their semantics.

We propose a paradigm shift with a lightweight security monitor:

- Users of vulnerable websites are automatically protected against a large class of attacks.
- Protocol specifications can be written and verified once, and then uniformly enforced.

WPSE Design and Implementation

The Web Protocol Security Enforcer (WPSE) is the first browser-side security monitor addressing the peculiar challenges of web protocols. The current prototype is implemented as an extension for Google Chrome.

Key Ideas

- Protocol Flow: WPSE describes web protocols in terms of the HTTP(S) exchanges observed by the web browser. The specification of the protocol flow defines the syntactic structure and the expected (sequential) order of the HTTP(S) messages. If protocol messages are not in the correct order or integrity constraints on messages are not satisfied, they are blocked and the automaton is reset to the initial state.
- Security Policies: Secrets in incoming messages are substituted with random placeholders before they enter the DOM and placeholders in outgoing requests are replaced with secrets only if sent to origins entitled to learn them.

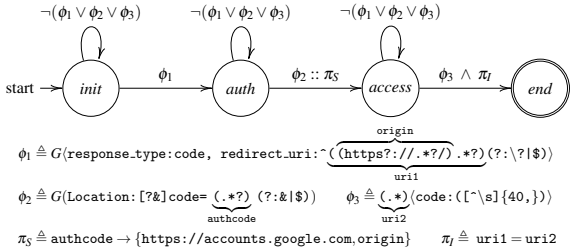


Figure 1: Automaton for OAuth 2.0 (authorization code mode)

Restrictions

- WPSE provides a significant improvement in security over standard web browsers but requires the specification of a protocol flow and a security policy.
- The implementation of the secrecy policies of WPSE is robust, but might break the website functionality if a trusted script needs to compute over a secret value exchanged in the protocol.
- The current prototype of WPSE suffers from some limitations due to the Google Chrome extension APIs.

Future Work

- Support for additional protocols e.g., e-payments
- Automatic techniques to synthesize WPSE policies from protocol specifications / browser traffic
- Embed WPSE into real browsers

Fortifying Web Protocols with WPSE

- Protocol Flow Deviations: WPSE blocks these attacks (e.g., some variants of CSRF and session swapping) as deviation from the intended protocol flow.
- Secrecy Enforcement: WPSE can prevent attacks where sensitive information is unintentionally leaked (e.g., via the Referer header) with a secrecy policy which specifies the origins that are entitled to receive a secret.
- Integrity Violations: WPSE can prevent these attacks by enforcing browser-side integrity checks.

Formal Results

- (H1) The protocol fulfills safety property P with a benign webpage.
- (H2) WPSE allows only a subset of the I/O sequences performed by the browser in an honest protocol run.
- (H3) Secrets are not leaked and securely stored by the browser.

→ Protocol fulfills P with a compromised browser monitored by WPSE

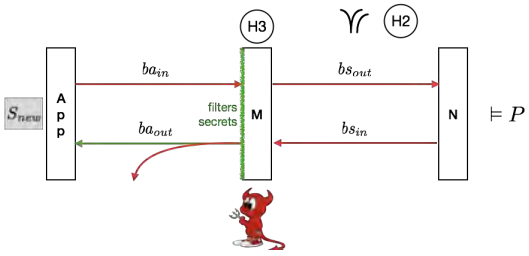


Figure 2: Visual description of Theorem 1

Experimental Evaluation

Manual investigation of 30 relying parties for each identity provider from Alexa top 100K. Analyzed both the authorization code mode and the implicit mode of OAuth 2.0.

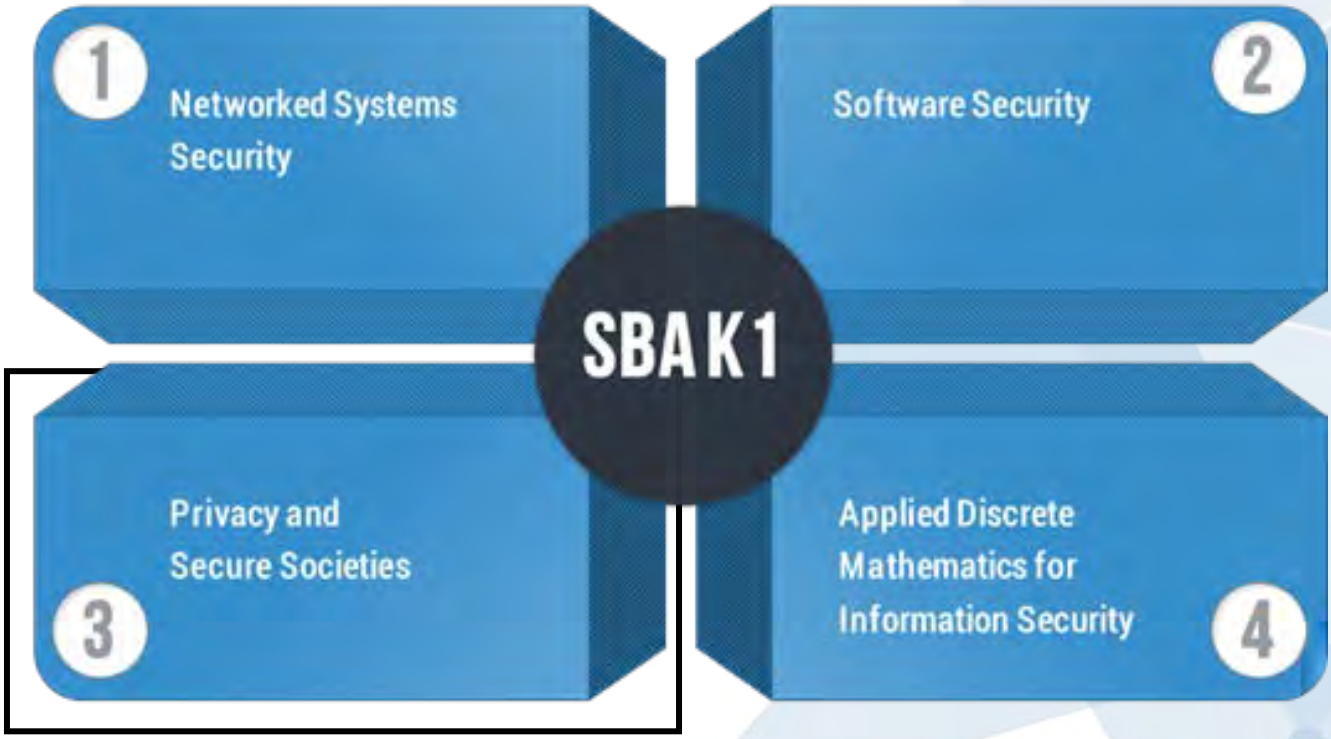
- Security: WPSE prevented the leakage of sensitive data on 4 different relying parties (advertisement libraries). 55 websites have been found affected by the lack or misuse of the state parameter.
- Compatibility: WPSE did not impact the browser performance or the time required to load webpages in a perceivable way. Problems due to security critical deviations in the protocol flow in 7 websites, e.g., authorization code is sent twice, second time over HTTP.

Stefano Calzavara, Riccardo Focardi, Matteo Maffei, Clara Schneidewind, Marco Squarcina, and Mauro Tempesta. WPSE: Fortifying web protocols via browser-side security monitoring. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 1493–1510, Baltimore, MD, August 2018. USENIX Association.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Privacy and Secure Societies

AREA 3



Block Me If You Can
A Large-Scale Study of Tracker-Blocking Tools

Georg Merzdovnik, Markus Huber, Damjan Buhov, Nick Nikiforakis, Sebastian Neuner, Martin Schmiedecker, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Third-party Services

- Websites and mobile applications rely on third-party services, like advertisements, analytics, social integration widgets, or CDN-residing JavaScript libraries.
- Benefits for developers are clear, but can have impact on users:
  - Increased tracking of users (third-parties are included in lots of different pages)
  - Direct attacks (like malware distribution through services)

Third-party Service Distribution

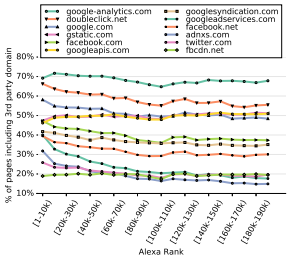


Figure 1: Distribution of the most popular third-party domains (TLD+1) in Alexa Top 200,000 websites in 10,000-intervals.

Third-party Blocking Analysis

- Various extensions exist to block third-party content
- Several Questions are still left open, e.g.:
  - What do they block?
  - What are their limits?

Browser Extension	Filter-Rules	Firefox Users	Chrome Users
AdBlock Plus (ABP)	ABP	18,689,656	10,000,000+
AdBlock	ABP	NA	10,000,000+
Ghostery	custom (proprietary)	1,337,831	2,348,209
uBlock (Origin)	ABP	1,243,409	3,852,990
AdBlock Edge	ABP	408,410	NA
Disconnect	custom (GPL)	265,773	797,097
Blur	custom (proprietary)	176,027	329,446
Privacy Badger	algorithmic	80,291	324,062

Figure 2: Common browser extensions to block online trackers, installations, and underlying filter rules (Aug. 2016).

Third-Party Reach

Webpages with Third-Party Inclusions per Company

	Desktop							*Mobile*			
	plain	AdBlock Plus	Disconnect	Ghostery	Privacy Badger	Ublock Origin	*all combined*	plain	EasyList	AdAway	MoaAB
Google	97	93	80	66	93	69	60	74	74	57	54
Facebook	47	44	5	2	4	39	0	6	6	6	6
Amazon	25	21	21	13	20	13	10	8	8	8	7
Twitter	24	21	6	1	19	19	1	1	1	1	1
Yahoo	18	6	4	2	3	2	1	14	14	14	0
AddThis	15	14	8	0	0	0	0	0	0	0	0
ComScor	14	10	1	0	1	0	0	2	2	0	0
AOL	11	0	1	0	1	0	0	0	0	0	0
Adobe	10	5	0	0	0	0	0	0	0	0	0
Quantcast	9	5	1	0	0	0	0	0	0	0	0
Conversant(ValueClick)	8	1	0	0	1	0	0	0	0	0	0
RadiumOne	6	1	0	0	0	0	0	0	0	0	0
Baidu	6	6	6	2	0	1	0	2	2	2	0
AudienceScience	5	0	0	0	0	0	0	0	0	0	0
Sizmek	5	0	0	0	1	0	0	0	0	0	0

Figure 3: Percentage of websites and Android applications reached by the Top 15 companies that provide third-party services. The results show the total reach (plain) as well as the reach after the application of each blocking solution.

Third-Party Inclusions not Blocked per Plugin

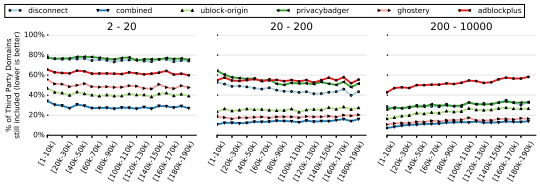


Figure 4: Sum of included third-party domains with 2-20/20-200/200-10,000 inclusions which are not blocked by a specific browser extension in relation to the plain profile. In all graphs: the lower an extension is on the y-axis, the better (i.e., less third-parties remaining).

Content-Blocking Capabilities

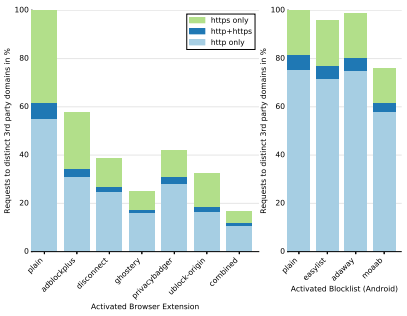


Figure 5: Protocols used for requests to distinct third-party domains.

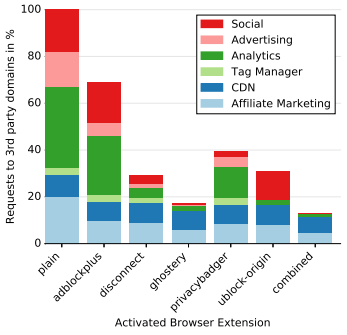


Figure 6: Categories blocked by different extensions, and all extensions combined. The data shows the categorized and aggregated numbers of the 30 most popular third-party services in our sample of 123,876 websites.

Conclusion

- A lot of traffic is still distributed through insecure channels (HTTP instead of HTTPS).
- Blocking tools differ greatly in performance and can have blind spots (e.g., not blocking smaller third parties).
- Third-party inclusions pose unique challenges on mobile devices (e.g., rooting of devices needed for blocking).

Georg Merzdovnik, Markus Huber, Damjan Buhov, Nick Nikiforakis, Sebastian Neuner, Martin Schmiedecker, and Edgar Weippl. Block me if you can: A large-scale study of tracker-blocking tools. In *IEEE European Symposium on Security and Privacy (EuroS&P)*, 2017.

This work was carried out within the scope of "u'smile", the Josef Ressel Center for User-Friendly Secure Mobile Environments, funded by the Christian Doppler Gesellschaft, A1 Telekom Austria AG, Drei-Banken-EDV GmbH, LG Nexera Business Solutions AG, NXP Semiconductors Austria GmbH, and Österreichische Staatsdruckerei GmbH. Furthermore, this work was supported by the Austrian Research Promotion Agency under grant 853264 (PriSAd), a joint project of Nimbusec GmbH and the St. Pölten University of Applied Sciences. For Stony Brook University, this research was supported by the National Science Foundation (NSF) under grants CNS-1527086 and CNS-1617593.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Blockchain Privacy
A Comparison of Mixing Techniques

Simin Ghesmati

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

- Blockchain addresses are **public keys** rather than **real identities**, different addresses could get **linked** together, and also linked to identities.
- Blockchain users can be deanonymized using different techniques such as **multi-input heuristic, transaction graph, side channel attacks, auxiliary information, or taint analysis** [1].
- Decentralized and centralized mixing techniques have been proposed to enhance blockchain privacy.
- The research questions are:
  - How resistant are these mixing techniques against privacy, e.g., theft, Sybil, or DoS attacks?
  - How do existing privacy techniques compare in terms of linkability, traceability, or transaction anonymity?
  - How could these techniques be aggregated or improved using new cryptographic techniques?

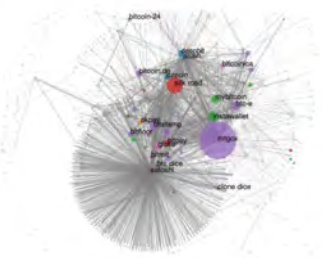


Figure 1: Bitcoin address clustering [2]

Mixing

In public blockchains such Bitcoin, transactions consist of multiple inputs and outputs which can be linked together. The mixing mechanism hides the links between inputs and outputs such that an attacker cannot link blockchain transactions. Various mixing techniques exist, and differ in terms of privacy and security levels. These techniques can be categorized into centralized and decentralized mixing.

Centralized Mixing

In centralized mixing, senders send their coins in equal amounts to a mixing service, which in turn, redistributes them to the intended recipients in a random permutation.

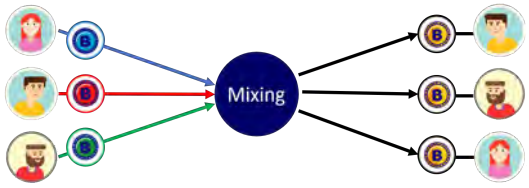


Figure 2: Centralized mixing service

Decentralized Mixing

Most of the decentralized mixing techniques are based on two ideas:

1. FairExchange

- Commitment transaction: commits a user to the coins exchange.
- Refund transaction: refunds a user's committed coins, in case the exchange protocol aborts.
- Claim transaction: allows a user to claim the other user's committed coins.

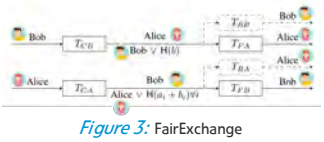


Figure 3: FairExchange

2. Coinjoin

- Agreeing on a set of inputs to spend, and a set of outputs to pay to.
- Users separately sign a transaction.
- One of the users posts the transaction to the network.

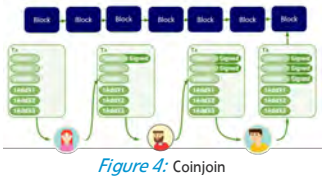


Figure 4: Coinjoin

Evaluation of Mixing Techniques

Techniques	Decentralized	Recipient privacy	Sender privacy	Payment value privacy	BTC Compatible	DoS resistance	Sybil resistance	Theft resistance	No interactivity with other users	Fair mixing fee	No interaction with the payee	Number of Trx	Min mixing time (Block)
Mixing website/services	○	○	◑	◑	●	●	●	○	●	○	●	2	2
CoinSwap	○	○	◑	○	●	●	●	●	●		○	4	2
Mixcoin	○	○	◑	◑	●	●	●	◑	●	◑	●	2	2
Blindcoin	○	○	●	◑	●	●	●	◑	●	○	●	2	4
Blindly Signed Contracts	○	○	●	◑	◑	●	●	●	●	◑	○	4	3
TumbleBit	○	○	●	◑	●	●	●	●	●	●	○	2	2
Fairexchange	●	○	◑	○	●	●	●	●	○	●	●	4	
CoinJoin	●	○	◑	○	●	○	○	●	○	●	●	1	1
Coinshuffle	●	○	●	○	●	○	○	●	○	●	●	2	2
Xim	●	○	◑	○	●	●	●	●	○	◑	●	4	hours
Coinparty	●	○	●	○	●	○	○	◑	○	●	●	2	2
Coinshuffle++	●	○	●	○	●	◑	○	●	○	●	●	2	2
Valueshuffle	●	●	●	●	◑	◑	○	●	○	●	●	1	1

Conclusion

- Centralized techniques rely on third parties.
- Most of the decentralized techniques prevent theft, but suffer from DoS and Sybil attacks.
- Better privacy requires higher delays.
- Most of the proposed techniques are not widely used in the real world.
- Proper techniques are selected according to user requirements.

[1]Joseph Bonneau, Arvind Narayanan, Andrew Miller, Jeremy Clark, Joshua A Kroll, and Edward W Felten. Mixcoin: Anonymity for bitcoin with accountable mixes. In *International Conference on Financial Cryptography and Data Security*, pages 486–504. Springer, 2014.
[2]Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M Voelker, and Stefan Savage. A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 127–140, 2013.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Fingerprinting Relational Data Sets

Tanja Šarčević, Rudolf Mayer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Fingerprinting techniques, which can be seen as a personalized version of generic watermarks applied to a digital object, can be utilized as a mechanism enabling ownership attribution. They generally embed a pattern in the data, i.e., they distort the original data set to a certain extent. A good fingerprint should (i) be recognizable by the original owner of the data, (ii) not be detectable (and consequently, removable) by recipients of the data, (iii) be robust to intentional or unintentional modifications of the data, and (iv) not lower the utility of the data too much.

The type of data in the dataset can be the crucial point for evaluating fingerprinting scheme effectiveness. Categorical data are shown to give rise to more problems with embedding the fingerprint compared to numerical data, yet the appropriate fingerprinting scheme for categorical data is necessary; otherwise, the domain of fingerprinting applications is very limited.

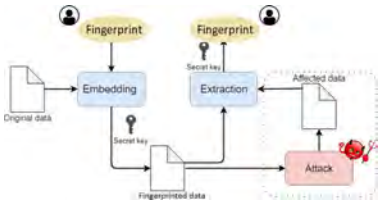


Figure 1: Fingerprinting workflow

Fingerprinting Numerical Data

- **AK Scheme**[2]: pseudo-random marking pattern
- **Block Scheme**[3]: binary image used as fingerprint information
- **Two-level Scheme**[4]: separate patterns for owner and the recipient

Utility Evaluation

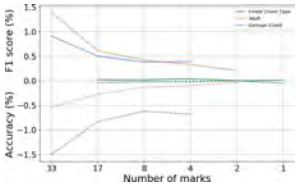


Figure 2: Classification performance of fingerprinted datasets

Data utility may be measured via its effect on machine learning model performance [5]. The representative results with Random Forest show rather small performance decreases, up to 1.5%. The performance drop is bigger for datasets with more introduced marks as well as for small datasets.

Data Utility Under Malicious Attacks

The attacks are additionally decreasing dataset's utility. The analysis shows the decrease in utility of 5 different classifiers under attacks. The results show that modifying data such that the fingerprint is not likely to be extracted anymore, the data loses on its utility significantly [6].

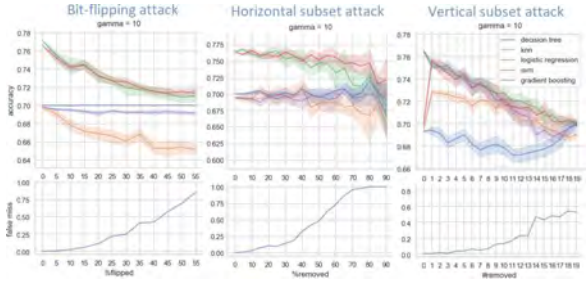


Figure 3: Data utility decrease by strengthening the attacks

Fingerprinting Categorical Data

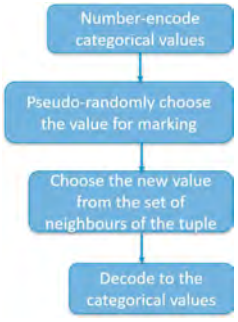


Figure 4: Fingerprinting scheme based on neighbourhood search

A novel scheme for fingerprinting categorical data in relational datasets is proposed in [1]. The scheme focuses on preserving the semantic relations between attributes, and thus limiting the perceptibility of marks, and the effects of the fingerprinting on the data quality and utility.

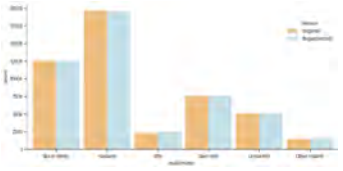


Figure 5: Distribution of a categorical attribute before and after fingerprinting

Robustness Evaluation

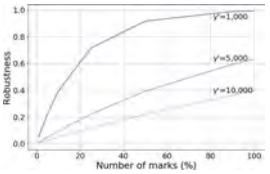


Figure 6: Additive attack

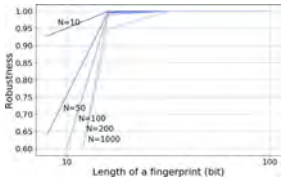


Figure 7: Misdiagnosis false hit

The schemes with less marks embedded in data are generally more susceptible to malicious attacks (actions on the dataset with the goal of removing the fingerprint). The main step for gaining robustness is choosing smaller fingerprint and embedding more marks.

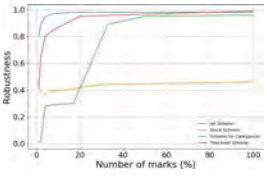


Figure 8: Subset attack

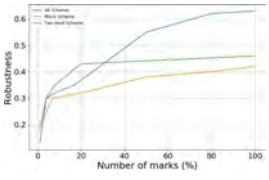


Figure 9: Bit-flipping attack

[1] Šarčević, T., Mayer, R.: A Correlation-Preserving Fingerprinting Technique for Categorical Data in Relational Databases. In 35th International Conference on ICT Systems Security and Privacy Protection (IFIP SEC 2020). Springer (2020).
[2] Li, Y., Swarup, V., Jajodia, S.: Fingerprinting relational databases: Schemes and specialities. IEEE Transactions on Dependable and Secure Computing 2(1), 34–45(2005).
[3] Liu, S., Wang, S., Deng, R.H., Shao, W.: A block oriented fingerprinting scheme in relational database. In: International Conference on Information Security and Cryptology. Springer (2004).
[4] Guo, F., Wang, J., and Li, D.: Fingerprinting Relational Databases. In ACM Symposium on Applied Computing (SAC). (2006).
[5] Šarčević, T., Mayer, R.: An Evaluation on Robustness and Utility of Fingerprinting Schemes. In Machine Learning and Knowledge Extraction: International Cross-Domain Conference (CD-MAKE). Springer (2019).
[6] Šarčević, T., Mayer, R.: Data Utility Assessment in Fingerprinted Datasets under Malicious Attacks. *Submitted for publication* (2020).

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

"I Have No Idea What I'm Doing" - On the Usability of Deploying HTTPS

Katharina Krombholz, Wilfried Mayer, Martin Schmiedecker, Edgar Weippl

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

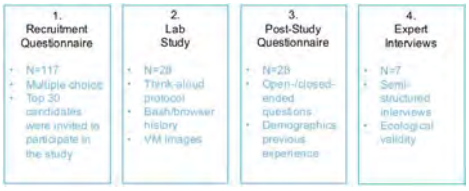
SBA Research

Problem & Motivation

- Explore **reasons for TLS misconfigurations**
- Usability from the **administrator's perspective**

- **Configure HTTPS on Apache**
  - Start with HTTP-configured Apache
  - Finish with secure HTTPS configuration

Methodology



Different parts of the study

- **Lab Study**
  - Get a certificate – Interact with CA
  - Change correct parameters – Harden configuration
  - Testing

Analyzed Results

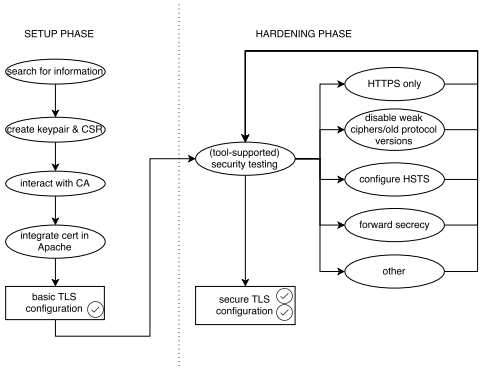
ID	Grade	Errors / Warnings / Highlights	Cipher Strength Score	Key Exchange Score	Protocol Support Score	Common Name	Key Size	Certificate Chain Length	Used Provided CA to Sign Encrypted Private Key	SSL 2	SSL 3	TLS 1.0	TLS 1.1	TLS 1.2	RC4 Support	Vulnerable to POODLE (SSL 3)	Forward Secrecy	HSTS	HPKP
P1	A	2	90	90	95	web.local	4096	3	●	○	○	○	●	●	○	○	●	○	○
P2	B	3	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	●	○	○
P3	B	2,3	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	●	●	○
P4	A		90	90	95	web.local	2048	3	●	○	○	○	●	●	○	○	●	○	○
P5	B		90	90	95	web.local	4096	1	●	○	○	○	●	●	○	○	●	●	○
P6	B	3	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	●	○	○
P7	Not valid																		
P8	C	3-6,8	90	90	50	web.local	2048	1	●	●	○	○	●	●	◐	○	○		
P9	B	1-3	100	90	95	web.local	4096	1	●	○	○	○	●	●	○	○	●	●	●
P10	B	1-3	90	90	95	web.local	4096	1	●	○	○	○	●	●	○	○	●	●	●
P11	B	3,4	90	90	95	web.local	2048	1	●	●	○	○	●	●	○	○	◐	○	○
P12	B	2,3	90	90	95	web.local	4096	1	●	○	○	○	●	○	○	○	●	●	○
P13	B	3	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	◐	○	○
P14	A-	4	90	90	100	raspberrypi	2048	1	○	○	○	○	○	●	○	○	◐	○	○
P15	C	4,7	50	90	95	-	2048	1	○	○	○	○	●	●	●	○	◐	○	○
P16	A-	4	90	90	95	web.local	2048	3	●	○	○	○	●	●	○	○	◐	○	○
P17	B	2,3	90	90	95	web.local	3096	1	●	○	○	○	●	●	○	○	●	●	○
P18	Not valid																		
P19	B	2,3	90	90	95	web.local	2048	1	●	●	○	○	●	●	○	○	●	○	○
P20	B	2,3	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	●	●	○
P21	B	3,4	90	90	95	Test	2048	1	●	○	○	○	●	●	○	○	◐	○	○
P22	B	3,4	90	90	95	web.local	2048	1	●	○	○	○	●	●	○	○	◐	○	○
P23	Not valid																		
P24	A	2	90	90	97	web.local	2048	3	●	○	○	○	●	●	○	○	●	●	○
P25	B	3	90	90	95	SME	4096	1	●	○	○	●	●	●	○	○	◐	○	○
P26	Not valid																		
P27	B	3,4	90	90	95	web.local	4096	1	●	○	○	○	●	●	○	○	◐	○	○
P28	A	2	90	90	95	web.local	4096	3	●	○	○	○	●	●	○	○	●	●	○

Analyzed Apache configurations

Statements

- "It seems that there is already a certificate called snakeoil, why can't I use this one?" (P7)
- "I have absolutely no idea what I'm doing. Neither am I aware of whether my online source is trustworthy." (P23)
- "The configuration process is fiddly and one has to google tons of pages to get it right. Even then one cannot be sure to have a good configuration, because SSL vulnerabilities are discovered almost on a regular basis." (P9)

Identified Workflow



Schematic representation of a successful workflow

Usability Challenges in TLS Deployment

- Searching for information and finding the right workflow
- Creating a Certificate Signing Request (CSR)
- Choosing the appropriate cipher suites
- Strict HTTPS
- Multiple configuration files
- Finding the right balance between security and compatibility

Conclusion

- Configuring TLS on Apache is a challenging task, even for experienced users and we should take this serious!
- Administrators struggle with important security decisions.
- Concerns are mainly driven by compatibility.
- It is hard to find reliable information sources.

Katharina Krombholz, Wilfried Mayer, Martin Schmiedecker, and Edgar R. Weippl. "I Have No Idea What I'm Doing" - On the Usability of Deploying HTTPS. In *26th USENIX Security Symposium (USENIX Security 2017)*, 8 2017.
Wilfried Mayer, Katharina Krombholz, Martin Schmiedecker, and Edgar R. Weippl. Securing the Internet, One HTTP 200 OK at a Time. *USENIX login, Winter 2017*, 42, 12 2017.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung
wirtschaftsagentur wien
Die Stadt Wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

"If HTTPS Were Secure, I Wouldn't Need 2FA"- End User and Administrator Mental Models of HTTPS

Alexandra Mai and Katharina Pfeffer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

HTTPS is the most widely used protocol to secure communication on the internet. However, there are still many websites which are not sufficiently secured since they either still use HTTP, or a weak HTTPS configuration.

To understand the reasons for these problems, we conducted the first qualitative mental model study of HTTPS which assesses end users' and administrators' perceptions of HTTPS.

User Study

- semi-structured interviews
- pilot studies (N=6)
- main study participants (N=30)
  - 12 administrators
  - 18 end-users
- post-hoc validity study (N=9)

Design & Procedure:

- 3 drawing tasks
- multiple rounds of inductive coding
  - 2 rounds open coding
  - selective coding
  - axial coding

Results

HTTPS

- Model

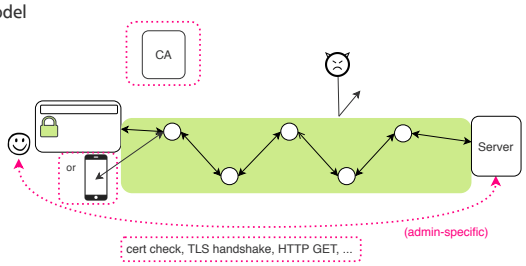


Figure 1: Model of HTTPS. Entities that are solely reflecting administrator mental models are highlighted (dashed boxed in pink).

- Anti-Model

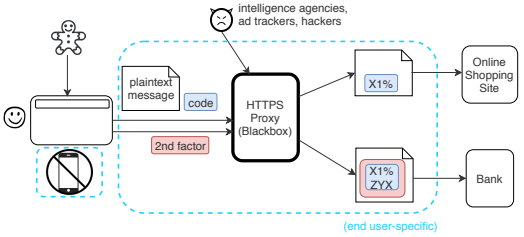


Figure 2: Anti-model of HTTPS. Entities that are solely reflecting end user mental models are highlighted (dashed boxes in blue).

Message - Encryption

- Model

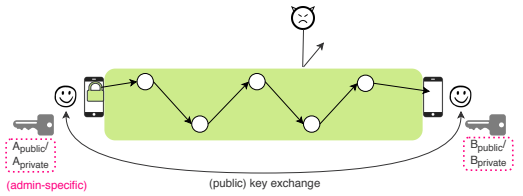


Figure 3: Model of message encryption. Entities that are solely reflecting administrator mental models are highlighted (dashed boxes in pink).

- Anti-Model

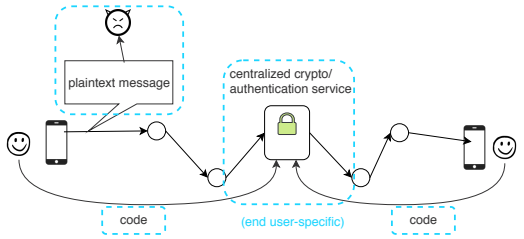


Figure 4: Anti-model of message encryption. Entities that are solely reflecting end user mental models are highlighted (dashed boxes in blue).

Conclusion

- Administrators: correct even if sparse mental models
- End users: conceptional and sparse mental models; sometimes wrong or non-existent
- Most participants mixed up authentication and encryption
- Incorrect mental models lead end users to mistrust the security of HTTPS and administrators to configure HTTPS incorrectly
- Recommendation: Provide encryption by default and highlight linkage between certificates and encryption during configuration

Katharina Krombholz, Karoline Busse, Katharina Pfeffer, Matthew Smith, and Emanuel von Zezschwitz. " if https were secure, i wouldn't need 2fa"-end user and administrator mental models of https. 2019.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

On the Usability of Authenticity Checks for Hardware Security Tokens

Alexandra Mai and Katharina Pfeffer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Problem & Motivation

Hardware security tokens (e.g., hardware wallets, Yubikeys) help users to keep stored secrets secure. However, recently reported attacks suggest that users cannot take the security guarantees of their devices for granted – even despite widely deployed authenticity checks.

Evaluate the effectiveness and usability of authenticity checks, we present (i) the first comprehensive market review analyzing authenticity checks of popular hardware security token and (ii) a large scale survey investigating user perceptions and usage of these checks.

Hardware Security Tokens

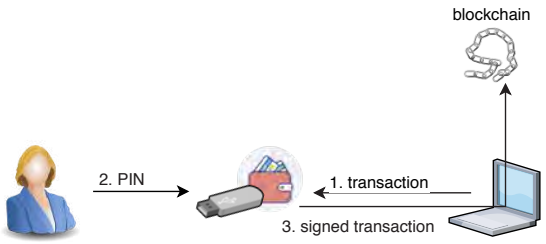


Figure 1: Simplified Hardware Wallet Transaction Model

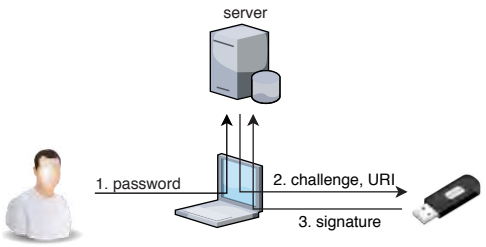


Figure 2: Simplified U2F Authentication Model

Market Review

- Reviewed tokens:
  - 4 YubiKey models
  - 5 hardware wallets
- Methodology: cognitive walkthroughs

		Hardware implants	Token replication	IC modification	Firmware modification	USB exploit	Token pre-initialization	Timing side-channels	Bus snooping	IC microprobing	Fault injection
		Hardware			Software			Secret Extraction			
Attestation / Countermeasure	Pkg	Tamper-evident	◐	◐	◐	◐	◐	◐	◐	◐	◐	◐
		Holographic sticker	○	○	○	○	○	○	○	○	○	○
	Hardware	Single-piece cast	●	◐	◐	○	○	○	○	◐	◐	○
		Openable device	◐	◐	○	○	○	○	○	○	○	○
		Secure element (co-processor)	○	●	●	○	○	○	●	◐	●	●
		Secure CPU	◐	●	●	○	○	○	●	●	●	●
	Software	Local firmware validation	○	○	○	●	○	○	○	○	○	○
		Remote firmware attestation	○	◐	○	●	○	○	○	○	○	○
		Key attestation	○	●	○	◐	○	○	○	○	○	○
		Manual firmware load	○	◐	○	●	○	◐	◐	◐	◐	◐

○ no prevention ● strong protection ◐ complicates attack/decreases usefulness

Table 1: Evaluation Framework: Mapping of Authenticity Checks to Attack Vectors

User Survey

- 2 discussion rounds with:
  - 9 HW security token users
  - 3 smartphone users
- Online questionnaire
  - 194 participants
  - 27–30 open/closed questions
  - 3 user groups: HW-Wallet, YubiKey, Smartphone

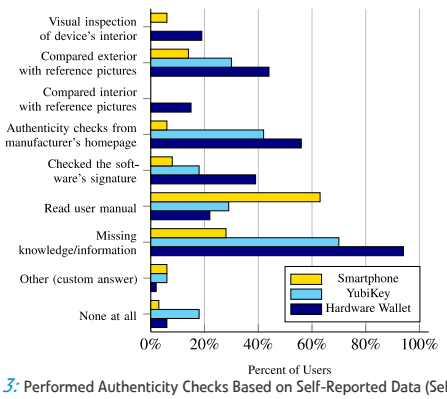


Figure 3: Performed Authenticity Checks Based on Self-Reported Data (Selection)

Conclusion

- Currently deployed authenticity checks—even in best-case implementations—are not sufficient to defeat all distribution attacks.
- Users incorrectly assess the existence and the security guarantees of many authenticity checks due to a lack of information and visibility.
- Recommendation: A combination of (i) secure CPUs or elements, (ii) remote firmware attestation, (iii) a recently proposed method for collaborative and verifiable key generation, and (iv) a user-friendly transparent design.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

RDA DMP Common Standard for Machine-actionable Data Management Plans

Tomasz Miksa

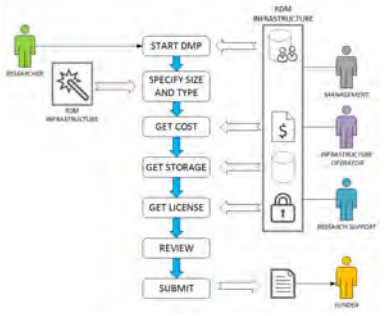
Problem & Motivation

Data Management Plans (DMPs) are documents accompanying research proposals and project outputs. They describe the data that is used and produced during the course of research activities, where the data will be archived, which licenses and constraints apply, and to whom credit should be given. The existing practice of writing DMPs is primarily driven by research funders who consider DMPs to be an awareness tool. However, DMPs are often perceived by researchers as an annoying administrative exercise that does not support data management activities.

We continue to need a human-readable narrative, but there is now widespread recognition that the DMP could have more thematic, **machine-actionable** richness with added value for all stakeholders. This includes **researchers, funders, repository managers, administrators, data stewards**, and so on. In short, everyone who is part of the larger ecosystem in which data is produced, transformed, exchanged, reused, and preserved. This added value can be created when parts of DMPs are pre-filled by **systems acting on behalf of stakeholders**. Similarly, information from DMPs can be used to trigger actions, for example, license and embargo selected by a researcher can be used to automatically fill out information on data deposited into a repository. Machine-actionability is also one of the main principles of the **European Open Science Cloud** (EOSC).

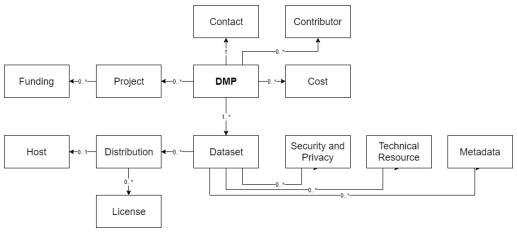
RDA DMP Common Standards working group

Research Data Alliance (RDA) recognized the importance of making DMPs machine-actionable and established the **DMP Common Standards working group to develop a standard** allowing for automatic exchange, integration, and validation of information provided in DMPs and facilitating the exchange of information between systems acting on behalf of stakeholders involved in the research life cycle, such as, researchers, funders, repository managers, ICT providers, librarians, etc.
The working group is chaired by Tomasz Miksa from SBA Research and has almost 200 members from all around the world. Within 18 months the group developed an application profile that acts as a standard for machine-actionable DMPs.



Common Standard for Machine-actionable DMPs

The **application profile** is meant for exchange of machine-actionable DMPs between systems. It is independent of any internal data organisation used by these systems. The application profile does not prescribe how information must be presented to the end user and does not enforce any specific logic on how this information must be collected or used. The application profile is an information carrier and the full machine-actionability can only be achieved when systems using the application profile implement appropriate logic.
This application profile is intended to cover a wide range of use cases and does not set any business (e.g. funder specific) requirements. It represents information over the whole DMP lifecycle, that is, it can express planned actions, as well as actions already performed.
The application profile is NOT intended to be a prescriptive template or a questionnaire, but to provide a re-usable way of representing machine-actionable information on themes covered by DMPs.



Adoptions of the common standard for machine-actionable DMPs

The application profile is the official output endorsed by Research Data Alliance. There are many adoptions pending at institutions from around the world:

- DMP Online by Digital Curation Centre (DCC) in the UK
- DMP Tool by California Digital Library (CDL) in the US
- DMP OPIDoR by Centre national de la recherche scientifique (CNRS) in France
- RDMO by Leibniz-Institut für Astrophysik Potsdam in Germany
- Data Stewardship Wizzard by Elixir research infrastructure in the EU
- Argos - OpenDMP by OpenAIRE and EUDAT research infrastructures in the EU
- F1000Research open research publisher in the UK
- Norwegian Open Research Data Infrastructure in Norway
- Haplo repository in the UK
- TU Wien, TU Graz, Uni Wien via FAIR Data Austria project

RESEARCH DATA ALLIANCE

Miksa, T., Walk, P., and Neish, P. (2019). RDA DMP Common Standard for Machine-actionable Data Management Plans. https://doi.org/10.15497/rda00039

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Review of the Stopp Corona App

Christian Kudera

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Background & Motivation

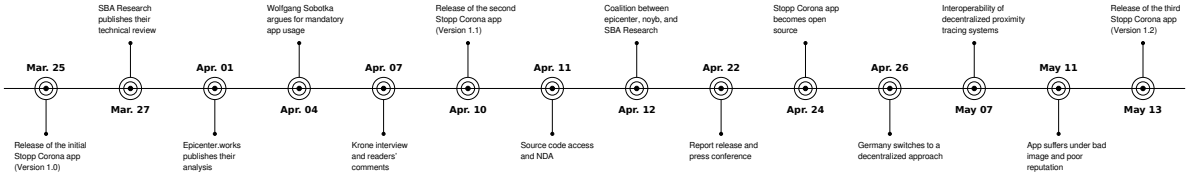
First "Corona-App" in the EU: Very early, the Austrian Red Cross published an initial version of a "Contact Tracing App" very early on March 25th. The app is operated by the Austrian Red Cross (a non-profit organization) and was developed by Accenture.

SBA Research already analyzed the first version of the app and reverse engineered its Android version for data security.

After a heated public debate – also about contact tracing in general –, the source code was submitted to the non-profit organizations epicenter.works, noyb.eu and SBA Research for a detailed technical review on data protection, IT security and legal aspects. Together, we published our findings and 26 recommendations in a detailed report. The results were presented at a press conference on April 22, 2020; more than half of the recommendations were fixed shortly afterwards.

We have not received any payment for this source code review. We see it as our task as an independent research institute to educate the public as neutrally as possible about the technical details.

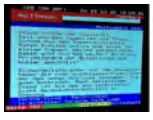
Timeline



Media Coverage (Version 1.0)

"From our point of view, the Android version of the Stopp Corona app adequately fulfils the requirements of data economy and respect for data protection in relation to the objectives of the app. Only the tracking of the handshakes should continue to be monitored critically, […]"

Markus Klemen, CEO of SBA Research

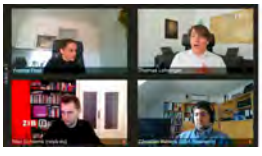


Media Coverage: Live Ticker (derstandard.at) (Mar 27) diepresse.com (Mar 27) futurezone.at (Mar 27) krone.at (Mar 27) ots.at (Mar 27) derstandard.at (Mar 27) vol.at (Mar 27) help.orf.at (March 30) Interview krone.at (Apr 07) news.orf.at (Apr 09)

Media Coverage (Version 1.1)

"A statistics function was built into the app, which transmitted the exchange of contacts via Bluetooth and the receipt of infection messages to the Red Cross. The statistics function was removed immediately due to our urgent recommendation."

Christian Kudera, IT Security Expert of SBA Research



Media Coverage: Ö1 Mittagsjournal, derstandard.at, heise.de, futurezone.at, ZIB 13:00, Computerwelt, PA Rotes Kreuz, diepresse.com, derbrutkasten.com, Ö3, NÖN.at, krone.at, OÖNachrichten (all Apr 22) FM4 Morning Show (Apr 23)

Source Code Review (Version 1.1)

Findings

- No critical security vulnerabilities
- 26 recommendations
  - 9 technical, 17 legal
    - 16 fixed in a hotfix
    - Three were fixed in version 1.2.
    - Four require a change in architecture.
    - Three legal recommendations will not be addressed.

Impact

- The source code was released to the public according to our recommendation.
- Press Release of the Austrian Red Cross after publishing our analysis:
  - "Constructive feedback from academia and civil society brings further improvements to the app and strengthens trust"

Figure 1: 50-page report published by SBA Research, epicenter.works, and noyb.eu

Figure 2: The source code was released to the public according to our recommendation

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Synthetic Data
Privacy Evaluation and Disclosure Risk Estimates

Markus Hittmeir, Andreas Ekelhart, Rudolf Mayer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Relation between Data Utility and Privacy

The poster on the *utility of synthetic data for machine learning* concerned three synthetic data generation tools: the Synthetic Data Vault (**SDV**), the DataSynthesizer (**DS**), and synthpop (**SP**). We obtained two main results:

- **SP** with standard settings tends to achieve better utility scores than SDV.
- For the **DS**, the results vary depending on the amount of noise injected by differential privacy.

Synthetic data with larger differences to the original (see Figures 1 and 2) tends to perform worse on certain tasks. On the other hand, it may provide better protection of the sensitive information in the original dataset. We hence complemented our utility evaluation with a privacy analysis.

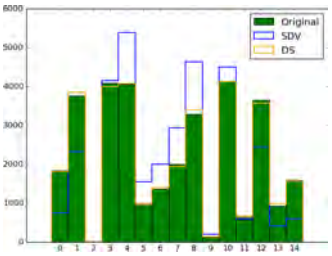
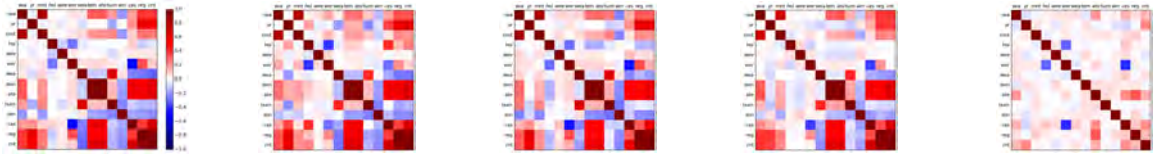


Figure 1: Distributions of the attribute 'occupation' on Adult Census data



(a) Real *(b)* DS *(c)* SP *(d)* SDV *(e)* DSP

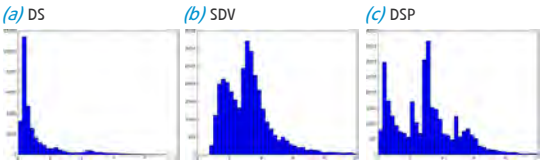
Figure 2: Heatmaps showing correlations on Bike Sharing data; red is direct, blue indirect correlation. **SDV** and **DSP** show larger differences to the original than **DS** and **SP**.

Similarity Analysis and Distance Measures

For the assessement of privacy provided by the synthetic datasets, we were interested in similarities between original and synthetic data samples. For each row in the synthetic data, we hence computed the distance to the nearest neighboring sample in the real data (=minimum distance).

General Findings

In Figure 3, we have the minimum distance on the x-axis and the number of samples on the y-axis. We see that **DS** constructs many samples that are similar to original ones, whereas **SDV** and **DSP** appear to provide better protection of the privacy of individuals in the real data. Note that the histogram for **SP** looks very much like the one of **DS**.



(a) DS *(b)* SDV *(c)* DSP

Figure 3: Minimum distances for the Adult Census dataset

We observed similar results on most of the datasets we considered. They appear to confirm that the better utility scores of **DS** and **SP** may be explained by the fact that these tools generate more synthetic samples that are close to original ones.

Disclosure Risk Estimates

In general, two possibilities of information disclosure are distinguished:

- **Identification disclosure** happens when a record in the dataset is linked to a certain individual.
- **Attribute disclosure** means that an attacker is able to infer someone's value of a sensitive attribute.

Our research focused on estimates for attribute disclosure risks. In Ref. 2, we generalized the Correct Attribution Probability (CAP) approach established in Ref. 1. This technique measures attribute disclosure risks for certain attack scenarios, consisting of:

1. A set of columns of the dataset, for which the attacker knows the values of their victim. Usually, these are quasi-identifying attributes like gender, age, or ZIP code.
2. A sensitive target column (e.g., health information or income).

We applied machine learning techniques to estimate the attacker's ability to retrieve the victim's value of target attribute.

General Findings

Table 1 shows the mean accuracy scores of one of said techniques in a certain scenario. Indeed, we observe lower risks on synthetic data than on the real data. However, the risks of **DS** and **SP** are higher then the risks of **SDV** and **DSP**.

	Real	DS	SP	SDV	DSP
Risk	51.7	46.8	48.8	39.0	45.2

Table 1: Attribute disclosure risk due to ENS from Scenario (1), Table 6 in Ref. 2.

Conclusion and Future Work

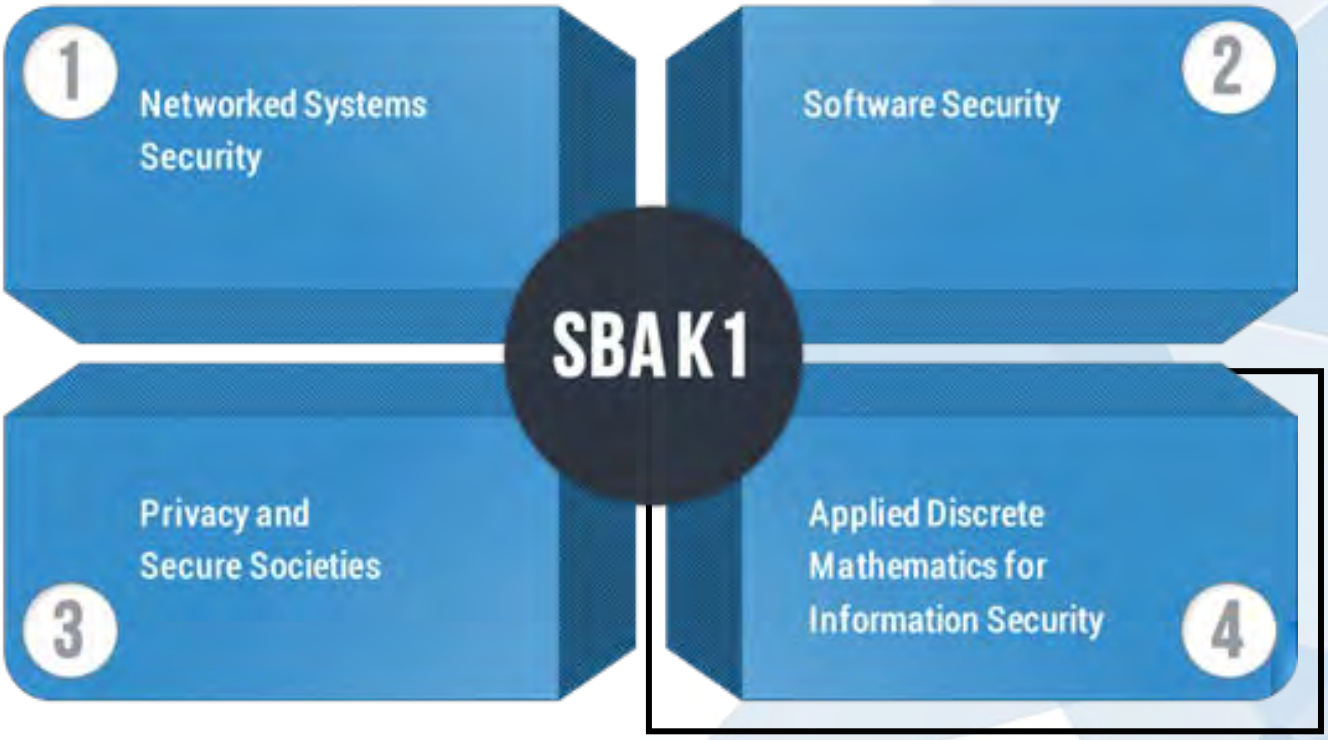
Synthetic data generation tools that performed better on utility tasks showed higher privacy and disclosure risks. All in all, we observed a trade-off between utility and privacy. As a consequence, one of our future goals is to optimize synthetic data for certain tasks and privacy requirements.

1. Taub et al., *Differential Correct Attribution Probability for Synthetic Data: An Exploration*, In: Proceedings of the 8th PSD, 2018.
2. M. Hittmeir, A. Ekelhart, R. Mayer, *A Baseline for Attribute Disclosure Risk in Synthetic Data*, In: Proceedings of the 20th CODASPY, 2020.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Applied Discrete Mathematics for Information Security

AREA 4



Algebraic Models for Covering Arrays

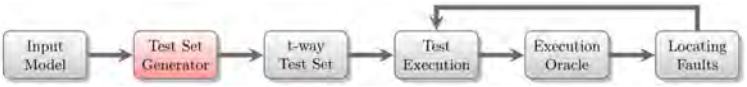
Dimitris E. Simos, Ludwig Kampel, Bernhard Garn, Ilias S. Kotsireas

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Covering Arrays and Optimal Covering Array Generation

Covering Arrays for Combinatorial Testing

- Covering Arrays (CAs) provide the theoretical means for Combinatorial Testing (CT).
- Columns of a CAs map to the parameters of a system under test.
- Rows of a CA encode the individual test cases.
- CAs guarantee that derived test sets **cover** all t**-way interactions**.
- Smaller test suites and reduced testing costs can be achieved by constructing CAs with less rows.



The Optimal Covering Array Generation Problem

- Given a **strength** t, a number of columns k and the respective column alphabet size v.
- Construct a covering array $\mathbf{CA}(N; t, k, v)$ with the smallest number of rows N.
- Direct constructions for optimal CAs, i.e., minimal N, exist only for boundary and some special cases.
- The general problem of actually constructing **optimal CAs** remains unsolved.

Algebraic Modelling of Covering Arrays

Definition of Covering Arrays

- Defining property of a $\mathbf{CA}(N; t, k, v)$:
  - For any sub-matrix comprised by t different columns it holds that
  - All $\{0, \ldots, v-1\}^t$ t-tuples appear at least once as a row

Covering Arrays as Solutions of Equation Systems

- By virtue of an appropriate algebraic structure, e.g.:
  - Integral domain R with unity
  - That has t linearly independent elements, when interpreted as $\mathbb{Z}_v$-module

Theorem. Let R be a ring and $(R, a_1, \ldots, a_t)$ have the v-ary t-way interaction distinguish property, and $X := (x_{i,j})$ be an $N \times k$ array of variables. Then any solution to the following system of equations in the unknowns $x_{i,j}$ yields a $\mathrm{CA}(N; t, k, v)$:

1. $\forall i \in \{1, \ldots, N\}, \forall j \in \{1, \ldots, k\}$

$$\prod_{r=0}^{v-1} (x_{i,j} - r) = 0. \tag{1}$$

2. $\forall C \in \binom{\{k\}}{t}$, $\forall (u_1, \ldots, u_t) \in [v]^t$:

$$\mathbf{prod}(X \cdot \iota_{t,k}^{C}(a_1, \ldots, a_t) - \mathbf{1} \cdot (u_1, \ldots, u_t) \cdot (a_1, \ldots, a_t)^T) = 0. \tag{2}$$

Algebraic and Symbolic Methods

$$\begin{pmatrix} 0 & 0 & x_1 \\ 1 & 0 & x_2 \\ 0 & 1 & x_3 \\ 1 & 1 & x_4 \end{pmatrix} \cdot \begin{pmatrix} a \\ 0 \\ b \end{pmatrix} = \begin{pmatrix} bx_1 \\ bx_2 + a \\ bx_3 \\ bx_4 + a \end{pmatrix}$$

$$\begin{aligned}
x_1(x_1-1) &= 0 \\
x_2(x_2-1) &= 0 \\
x_3(x_3-1) &= 0 \\
x_4(x_4-1) &= 0 \\
(bx_1-b)(bx_2+a-b)(bx_3-b)(bx_4+a-b) &= 0 \\
(bx_1-b)(bx_2-b)(bx_3+a-b)(bx_4+a-b) &= 0 \\
(bx_1-a-b)(bx_2-b)(bx_3-a-b)(bx_4-b) &= 0 \\
(bx_1-a-b)(bx_2-a-b)(bx_3-b)(bx_4-b) &= 0 \\
b^2x_1(bx_2+a)x_3(bx_4+a) &= 0 \\
b^2x_1x_2(bx_3+a)(bx_4+a) &= 0 \\
(bx_1-a)b^2x_2(bx_3-a)x_4 &= 0 \\
(bx_1-a)(bx_2-a)b^2x_3x_4 &= 0
\end{aligned}$$

```
S:=RationalField();
P:=PolynomialRing(S,k*N+2);
R:=PolynomialRing(S,k*N);
M := ZeroMatrix(P,k,N);
for i in [1..k] do
    for j in [1..N] do
        M[i][j] := P.((i-1)*N+j);
    end for;
end for;
```

- Multiple approaches to solve the equation systems:
  - Algebraic Solvers (Gröbner Bases)
  - SAT-Solvers
  - Search techniques and high performance computing

Algorithmic Formulation for Covering Arrays

- The algebraic model allows to formulate an algorithm for CA computation.
- Above theorem yields the following algorithm.

Algorithm 1 AlgebraicSearchCAs

```
INPUT: N, t, k, v
Require: t ≤ k
Create a symbolic N × k array X containing variables x_1, ..., x_Nk
EQall := ∅
for C ∈ binom({k}, t) do                          ▷ Add coverage equations
    for u ∈ [v]^t do
        EQ := prod(X · ι^C_{t,k}(a_1, ..., a_t) − 1 · (u_1, ..., u_t) · (v^0, ..., v^{t−1})^T) = 0
        add EQ to EQall
    end for
end for
for i = 1, ..., Nk do                            ▷ Add domain equations
    EQ := ∏_{j=0}^{v−1} (x_i − j) = 0
    add EQ to EQall
end for
Interpret EQall as subset of Q[x_1, ..., x_Nk]
V = Solve(EQall)                                  ▷ Call external solver
if V ≠ ∅ then
    return V;
else print "No CA exists";
end if
```

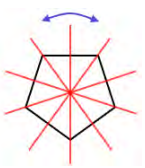
Computational Results

- The approach is exact and complete in nature.
- Full enumeration of **all optimal CAs** for specific parameters.

CA instance	Solver	# Vars	# Sols	CAN
$CA(4; 2, 3, 2)$	GB	12	48	4
$CA(5; 2, 3, 2)$	GB	15	1440	4
$CA(5; 2, 4, 2)$	GB	20	1920	5
$CA(8; 3, 4, 2)$	GB	32	80640	8
$CA(9; 2, 3, 3)$	C/MPI	27	$\geq 3 \cdot 10^6$	5

Future Work

- Enhancement of existing work by structuring equations.
- Symmetry breaking during solving process.
- Algebraic modellings of related designs.
- Hybridization with other approaches.



Bernhard Garn and Dimitris E. Simos. Algebraic modelling of covering arrays. In Ilias S. Kotsireas and Edgar Martínez-Moro, editors, *Applications of Computer Algebra*, pages 149–170, Cham, 2017. Springer International Publishing.
Bernhard Garn and Dimitris E. Simos. Algebraic techniques for covering arrays and related structures. *Electronic Notes in Discrete Mathematics*, 70:49 – 54, 2018. TCDM 2018 – 2nd IMA Conference on Theoretical and Computational Discrete Mathematics, University of Derby.
Ludwig Kampel, Dimitris E. Simos, Bernhard Garn, Ilias S. Kotsireas, and Evgeny Zhereshchin. Algebraic models for arbitrary strength covering arrays over v-ary alphabets. In Miroslav Ćirić, Manfred Droste, and Jean-Éric Pin, editors, *Algebraic Informatics*, pages 177–189, Cham, 2019. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Coverage and Distance Measurements of Test Sets

Dimitris E. Simos, Manuel Leithner, Rick Kuhn, Raghu Kacker

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Key Facts

Use Cases

- Measure coverage of existing test sets.
- Verify coverage of constructed Covering Arrays.
- Qualitative comparison of test sets with identical coverage with additional distribution and distance analysis.

Architecture

- Rust implementation:
  - ⇒ Native Executable (Linux, Windows, macOS)
  - ⇒ WebAssembly
- Modular input parsing.
- Machine- or human-readable output.

Implementation Highlights of CAmetrics

- Significantly fast based on experiments:
  - ⇒ Multi-threaded
  - ⇒ Multiple algorithms
- Low memory usage.
- Web UI and command line interface.
- Sophisticated constraint support.

Future improvements

- HPC and cluster implementations.
- Faster constraint processing.
- α-balance.

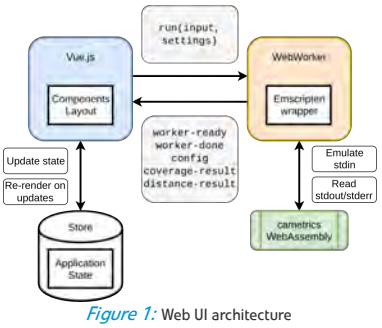


Figure 1: Web UI architecture

Coverage

- Simple t-way combination coverage: How many t-selections of parameters are fully covered?
- Simple $(t+1)$-way coverage: $t + \frac{\text{Covered } (t+1)\text{-tuples}}{\text{Total } (t+1)\text{-tuples}}$

Figure 2: Simple t-way combination and tuple coverage

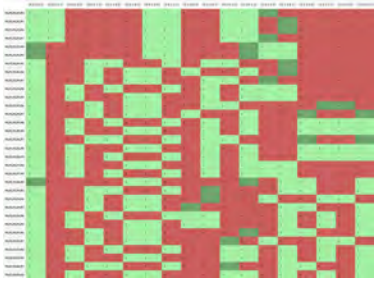


Figure 3: Coverage Map showing t-tuple coverage per parameter selection

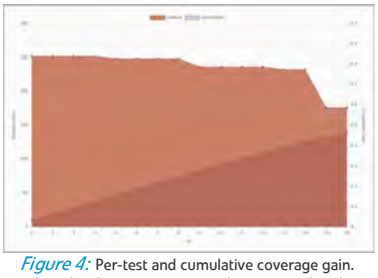


Figure 4: Per-test and cumulative coverage gain.
The coverage gain describes how many previously uncovered tuples are covered by a single test.

Select between Two Algorithms

Input: $N \times k$ array, alphabet size v, strength t

`cametrics-fast`

- Great general-purpose performance
- $\mathcal{O}(\binom{k}{t} v^t)$ memory
- $\mathcal{O}(\binom{k}{t} N)$ time

`cametrics-light`

- Prevents memory explosion, great for binary/small arrays
- $\mathcal{O}(\binom{k}{t})$ memory
- $\mathcal{O}(\binom{k}{t} N v^t)$ time

Distance Metrics

Inter-test distance: baseline for success

- (Generalized) Hamming Distance: How many parameter values differ between tests?
- Total Cartesian/Euclidian Distance of test T to array A:

$$CD(A,T) = \sum_{n=1}^{|A|} \sqrt{\sum_{i=1}^{k} (A_{ni} - T_i)^2}$$

Balance is everything!

- *Balanced* array: Each parameter value (or t-tuple) appears roughly the same number of times.

Modified χ^2 Distance:

- How close to ideal distribution of parameter values?

Ideal distribution: $D'_{ij} = \frac{1}{v_i} | i \in \{1, \ldots, k\}, j \in V_i$
Actual distribution:

$$D_{ij} = \frac{1}{N} \sum_{n=1}^{N} \begin{cases} 1 & \text{if } A_{ni} = V_{ij} \\ 0 & \text{otherwise.} \end{cases}$$

Modified χ^2 Distance: $(\sum_{i=1}^{k} \frac{v_i - 1}{v_i + 1}) - \frac{\sum_{i=1}^{k} \sum_{j=1}^{v_i} \frac{(D_{ij} - D'_{ij})^2}{D_{ij} + D'_{ij}}}{2}$

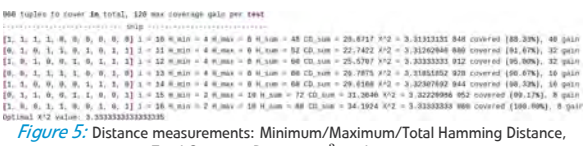


Figure 5: Distance measurements: Minimum/Maximum/Total Hamming Distance, Total Cartesian Distance, χ^2, and coverage

D Richard Kuhn, Itzel Dominguez Mendoza, Raghu N Kacker, and Yu Lei. Combinatorial coverage measurement concepts and applications. In *2013 IEEE Sixth International Conference on Software Testing, Verification and Validation Workshops*, pages 352–361. IEEE, 2013.
Manuel Leithner, Kristoffer Kleine, and Dimitris E Simos. CAmetrics: A tool for advanced combinatorial analysis and measurement of test sets. In *2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 318–327. IEEE, 2018.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Fault Localization for Web Security Testing

Dimitris E. Simos, Bernhard Garn, Manuel Leithner, Yu Lei, Angelo Gargantini

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research
MATRIS

Combinatorial Testing & Combinatorial Fault Analysis

Motivation

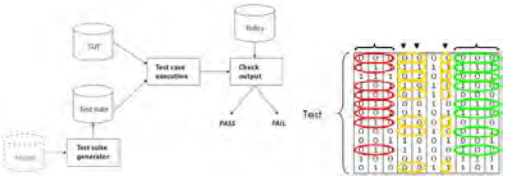
- We cannot test everything.
- Exhaustive search of problem space increases time needed exponentially.
- Automated detection of security vulnerabilities.

Combinatorial Security Testing (CST)

- Parameters and values provide abstract models of attacks
- Generated test sets provide 100% coverage of t-way parameter value combinations.
- Automated test set generation, execution and evaluation via dedicated test oracle.

Technical Challenges

- Generation of minimal t-way test sets is a hard combinatorial optimization problem.
- Modelling of parameters, values and constraints is domain-specific.
- Deploy CST to all application layers of information security.



Cross-Site Scripting (XSS)

- **Vulnerability**: Response from web server contains parts of unsufficiently sanitized user input.
- **Threat**: Attacker can execute malicious JavaScript.
- **Goal**: Automatically generate XSS attack vectors for testing purposes.
- **Targets**: HTTP parameters of web applications.
- **Model**: Parameters map to parts of the URL.
- **Example**: `<scr<script> ``> onLoad(;\>.`

Combinatorial Analysis of XSS Vulnerabilities

- XSS attack vectors generated with CST.
- Successful vectors are analyzed for their combinatorial structure
- Identification of XSS-inducing combinations provides insights.
- Approach evaluated against four sanitization functions from the Web Application Vulnerability Scanner Evaluation Project (WAVSEP).
- Results show effective identification of XSS-inducing combinations.

JSO	WS1	INT	WS2	EVH	WS3	PAY	WS4	PAS	WS5	JSE
"><script>	␣	';	␣	onError=	␣	alert(1)	␣	'>	␣	\>
"><script>	␣	'>	␣	onError=	␣	alert(1)	␣	'>	␣	\>
"><script>	␣	'>	␣	onError=	␣	alert(1)	␣	')	␣	\>
"><script>	␣	€	␣	onError=	␣	alert(1)	␣	'>	␣	\>
"><script>	␣	€	␣	onError=	␣	alert(1)	␣	')	␣	\>
"><script>	␣	'>>	␣	onError=	␣	alert(1)	␣	')	␣	\>
"><script>	␣	';	␣	onError=	␣	src="invalid"	␣	'>	␣	\>
"><script>	␣	'>	␣	onError=	␣	src="invalid"	␣	'>	␣	\>
"><script>	␣	'>	␣	onError=	␣	src="invalid"	␣	')	␣	\>
"><script>	␣	€	␣	onError=	␣	src="invalid"	␣	'>	␣	\>

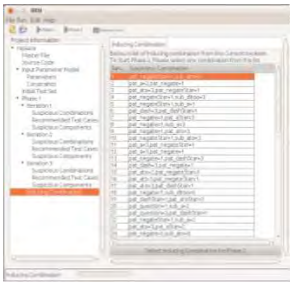
Combinatorial Fault Analysis (FLA)

Theory

- A combination is called suspicious if it appears only in a failing test case.
- A combination c is failure-inducing if any test f in which c is contained, fails.
- Identification of minimal failure-inducing combinations.
- Active research area in CT.

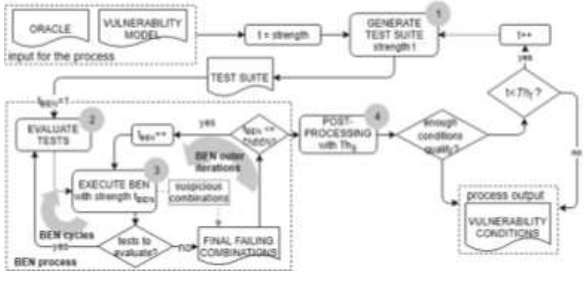
BEN Tool

- CT-based fault analysis tool.
- Input: executed t-way test set with pass/fail assignments.
- Output: ranking of combinations in terms of their likelihood to be failure-inducing.
- Adaptive approach: small number of additional tests might required.
- Written in Java and provides both GUI and CLI interfaces.



Fault-driven Combinatorial Process for Model Evolution in XSS

- Knowledge base (KB) contais model for XSS.
- Iterative evolvement of KB for XSS security testing of web applications.
- KB gives rise to attack strings for exploiting XSS vulnerabilities.
- Testing results are annoted and added back to KB.
- Process uses BEN tool internally.
- Increases capabilites of KB for subsequently requested attack models.



Bernhard Garn, Marco Radavelli, Angelo Gargantini, Manuel Leithner, and Dimitris E. Simos. A fault-driven combinatorial process for model evolution in xss vulnerability detection. In Franz Wotawa, Gerhard Friedrich, Ingo Pill, Roxane Koitz-Hristov, and Moonis Ali, editors, *Advances and Trends in Artificial Intelligence. From Theory to Practice*, pages 207–215, 2019.
Dimitris E. Simos, Kristoffer Kleine, Laleh Shikh Gholamhossein Ghandehari, Bernhard Garn, and Yu Lei. A combinatorial approach to analyzing cross-site scripting (xss) vulnerabilities in web application security testing. In Franz Wotawa, Mihai Nica, and Natalia Kushik, editors, *Testing Software and Systems*, pages 70–85, 2016.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Testing Methods for Composed Systems

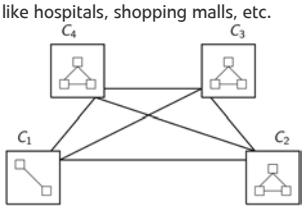
Dimitris E. Simos, Ludwig Kampel, Bernhard, Garn, Murat Ozcan

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Combinatorial Methods for Modelling Composed Systems

Complex systems composed of multiple components
Emerging in vast numbers in modern information society:

- Contemporary software design relies on **modular software architecture**, making it better understandable and maintainable
- A **modern vehicle** is a composed complex system in itself
- **Communicating autonomous vehicles** are even more complex
- **Smart buildings** like hospitals, shopping malls, etc.

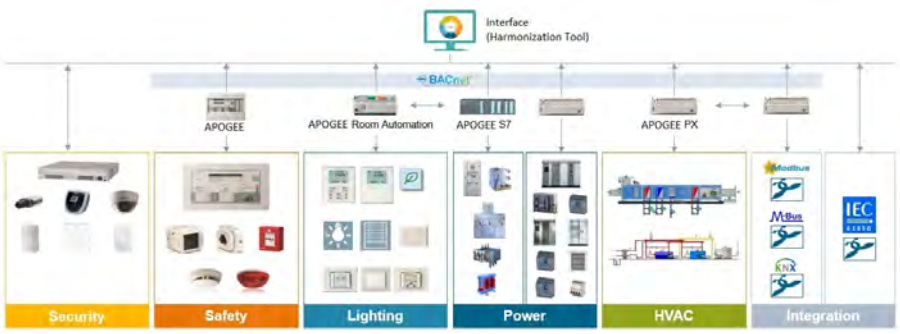


- Composed SUT with components C_1, C_2, C_3 and C_4, e.g., a car:
  - C_1: Wheels modelled via `producer`, `type`
  - C_2: Engine modelled via `fuel`, `drive-mode`, `filter`
  - C_3: Infotainment modelled via `streaming`, `audio`, `remote`
  - C_4: Communication modelled via `ip-version`, `connection`, `speed`

Methodology
Different in nature and in their application domains, these systems under test (SUTs) are **systems composed of sub-systems**:

- Devise a combinatorial model and a test suite for each sub-system
- Devise a combinatorial model for the unifying meta-system
- Apply a combinatorial construction to merge test suites of the components to a test suite for the whole SUT
- From theory we know: The **coverage of all t-way interactions is inherited** to the overall test suite

BACnet (Building Automation and Control networking) Protocol



- BACnet enables devices access via the network.
- Interoperability among different vendors' equipment.
- One operator interface to handle any device in the network.
- US, EU and ISO standard.

Combinatorial Methods for Testing BACnet

- BACnet models devices using an object oriented structure.
- Event Enrollment Objects (EEOs) provide an interface to communicate with devices.
- 5 million ways to configure an EEO.

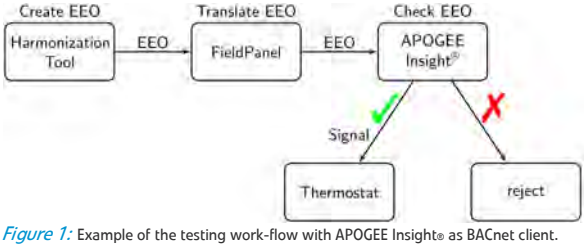


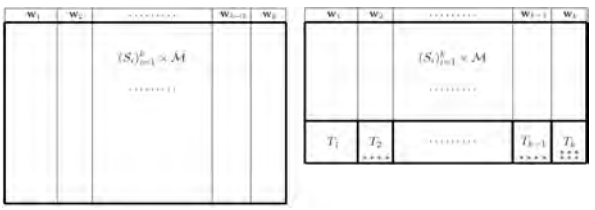
Figure 1: Example of the testing work-flow with APOGEE Insight® as BACnet client.

- (Nested) IPM for EEOs.
- Optimized test suite, maching the constraints of the application.
- Execute tests.

A Plug-in Construction for CAs Reflecting Composed Systems

- **Goal:** Construct CAs with more factors from CAs with less factors.
- **Idea:** Adapt plug-in construction from classic design theory for CAs.
- **Methodology:** Make use of coverage inheritance.
- **Application:** Combinatorial Testing for composed (Software) Systems.

Theorem. *Given an MCA* $\mathcal{M} = \mathsf{MCA}(N; t, k, (u_1, \ldots, u_k))$ *and two families* $T_i = \mathsf{MCA}(v_i; t_i, g_i, \mathbf{w}_i = (w_{i,1}, \ldots w_{i,g_i}))$ *and* $S_i = \mathsf{MCA}(u_i; t_i - 1, g_i, \mathbf{w}_i = (w_{i,1}, \ldots w_{i,g_i}))$ *of MCAs, for* $i = 1, \ldots, k$. *Then a* $\mathsf{MCA}(M; \tau, \sum_i g_i, (\mathbf{w}_1, \ldots, \mathbf{w}_k))$ *can be constructed, where* $M = N + \max_{i \in \{1,\ldots,k\}}\{v_i\}$ *and* $\tau = \min_{i \in \{1,\ldots,k\}}\{t_i, t\}$.



L. Kampel, B. Garn, and D. E. Simos. Combinatorial methods for modelling composed software systems. In *2017 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 229–238, 2017.
Dimitris E. Simos, Ludwig Kampel, and Murat Ozcan. Combinatorial methods for testing communication protocols in smart cities. In Roberto Battiti, Mauro Brunato, Ilias Kotsireas, and Panos M. Pardalos, editors, *Learning and Intelligent Optimization*, pages 437–440, Cham, 2019. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Testing Methods for SQL Injections

Dimitris E. Simos, Jovan Zivanovic, Manuel Leithner, Bernhard Garn

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research
MATRIS

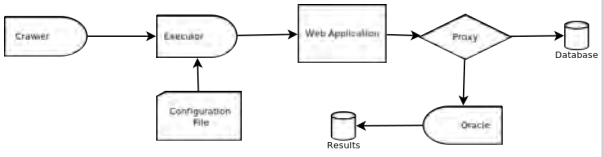
SQL Injection

- Unsanitized user input used inside *SQL* statements.
- Malicious user can execute arbitrary commands on the database.

SQL Injection Example

SQLInjector

- Developed prototype tool for executing *CT*-generated attack vectors.
- Uses a *database proxy* for collecting queries.
- Potentially malicious queries are compared against known valids.
- Changes in syntax indicate successful *SQL injection*.



Combinatorial Security Testing

- Grammar used to define the structure attack vectors with discretized parameters.
- Every parameter represents a parth that has a specific purpose inside the attack vector, e.g., quotation marks for escaping.
- *Input Parameter Model (IPM)* used to define the attack grammar.
- Each row of a *Covering Array (CA)* represents an attack vector.

Attack Model

```
InQl(5) ::= " | ' | %22 | %27 | ε
WhSpl(9) ::=  | /**/ | %20 | %09 | %A0 | %0A |
 %0B | %0C | %0D
Par(3) ::= ) | %29 | ε
Comml(9) ::= -- | # | /* | %00 | %23 | %2F |
 %2A | %2F%2A | ε
InVa(6) ::= 0 | 1 | 2 | a | b | c
CndV1(6) ::= 0 | 1 | 2 | true | !!true | !false
CndV2(6) ::= 0 | 1 | 2 | true | !!true | !false
Cmd(14) ::= OR | Or | oR | or | || | AND | ANd |
AnD | aND | anD | And | aNd | and | &&
AtkVec1(1) ::= InVa InQl Par WhSpl Cmd WhSpl
CndV1 WhSpl = WhSpl CndV2 WhSpl Comml WhSpl
```

1, 3, 1, 1, 1, 1, 0, 0	1%27/**/||/**/(Select/**/1+1)/**/#/**/
2, 2, 2, 2, 2, 2, 1, 0	2%22%20&&%20(sElEcT%202|2)%20#%20
0, 3, 0, 3, 0, 0, 2, 0	0%27 and (select 0*0) #

Case Study

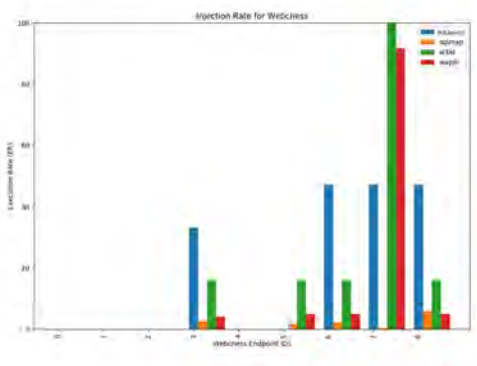
- **WAVSEP**: Well-known verification framework with known vulnerabilities, used to evaluate automated web application vulnerability scanners.
- **Webchess**: A self-hosted PHP-based online chess application.
- **geccBBlite**: Minimalistic bulletin board software.
- **OpenSchool**: Online School Management Interface.
- **zeusCart**: Open-source shopping cart for online stores.
- **uHotelbooking**: Online hotel reservation system.
- **modesecurity**: Web Application Firewall.

Attack vector comparison

- **sqlmap**: Well known automated SQL injection vulnerability scanner.
- **wapiti**: Web application vulnerability scanner written in *python*.
- **w3af**: Also an automated web application scanner written in *python*.

Evaluation

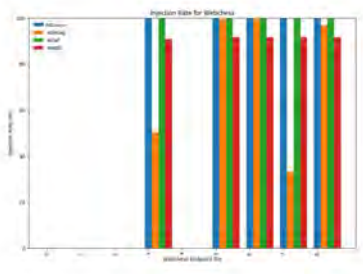
Execution rate for webchess against all vuln. scanners.



Tested parameters for each SUT

SUT	WAVSEP	WTF	webchess	geccBBlite	OpenSchool	zeusCart	uHotelbooking
Tested Params.	48	N/A	9	5	43	24	17
Vulnerable Params.	30	N/A	5	2	4	4	1

Injection Rate for webchess against all vuln. scanners.



Dimitris Simos, Jovan Zivanovic, and Manuel Leithner. Automated combinatorial testing for detecting sql vulnerabilities in web applications. In *14th International Workshop on Automation of Software Test (AST)*, 5 2019.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Testing Methods for XSS Vulnerabilities

Dimitris E. Simos, Bernhard Garn, Jovan Zivanovic, Manuel Leithner, Josip Bozic, Franz Wotawa

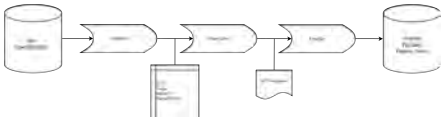
COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research
MATRIS

XSSInjections

- Unsanitized user input displayed on web application.
- Malicious users can add *HTML-Script* element.
- JavaScript gets executed in client machine.

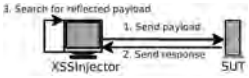
XSSInjector

- Prototype tool for executing CT generated attack vectors.
- Browser Oracle for checking executed JavaScript.



Reflection Oracle (RO)

- Verify whether the injected attack vector is present in response.
- False-positives and false-negatives easily possible.

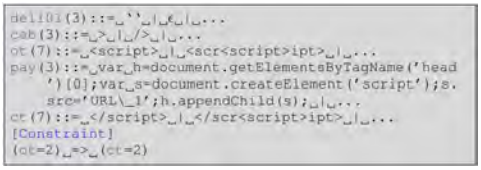


Browser Oracle (BO)

- Gets a request from executed attack vector.
- No false-positives.

Found Vulnerabilities

Attack Model

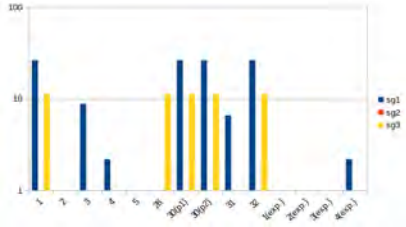


Case Study

- **WAVSEP**: Well known verification framework with known vulnerabilities, used to evalute automated web application vulnerability scanners.
- **Piwigo**: A PHP-based photo gallery software.
- **MyBB**: Self-hosted bulletin board application.
- **Koha**: An integrated library system written in Perl.
- **W3C tidy service**: Online tool for validating and fixing HTML code.

Evaluation

WAVSEP Injection Rage (Browser Oracle)



Number of total XSS injections for all SUT endpoints for both oracles.

SUT	t	Grammar 1		Grammar 2		Grammar 3	
		RO	BO	RO	BO	RO	BO
WAVSEP	2	347	57	316	0	70	5
	3	1125	229	891	2	198	19
Piwigo	2	48	20	115	5	37	4
	3	318	89	275	15	74	7
MyBB	2	46	4	51	0	24	0
	3	149	8	178	2	67	0
Koha	2	N/A	14	N/A	0	N/A	0
	3	N/A	29	N/A	0	N/A	1

Josip Bozic, Dimitris E. Simos, and Franz Wotawa. Attack pattern-based combinatorial testing. In *Proceedings of the 9th International Workshop on Automation of Software Test*, AST 2014, page 1–7, New York, NY, USA, 2014. Association for Computing Machinery.
Bernhard Garn, Marco Radavelli, Angelo Gargantini, Manuel Leithner, and Dimitris E. Simos. A fault-driven combinatorial process for model evolution in xss vulnerability detection. In Franz Wotawa, Gerhard Friedrich, Ingo Pill, Roxane Koitz-Hristov, and Moonis Ali, editors, *Advances and Trends in Artificial Intelligence. From Theory to Practice*, pages 207–215, Cham, 2019. Springer International Publishing.
D. E. Simos, B. Garn, J. Zivanovic, and M. Leithner. Practical combinatorial testing for xss detection using locally optimized attack models. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 122–130, 2019.
Dimitris E. Simos, Kristoffer Kleine, Laleh Shikh Gholamhossein Ghandehari, Bernhard Garn, and Yu Lei. A combinatorial approach to analyzing cross-site scripting (xss) vulnerabilities in web application security testing. In Franz Wotawa, Mihai Nica, and Natalia Kushik, editors, *Testing Software and Systems*, pages 70–85, Cham, 2016. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung seit 1999
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Combinatorial Testing of TLS, X.509 and IoT protocols

Dimitris E. Simos, Bernhard Garn, Manuel Leithner, Dominik Schreiber, Yu Lei, Franz Wotawa

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research
MATRIS

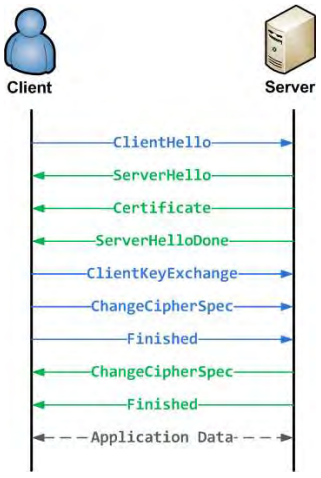
Transport Layer Security (TLS/SSL)

- Most common communications security protocol on the Internet.
- Provides confidentiality via symmetric encryption.
- Authenticity of servers provided via X.509 certificates:
  - Client authenticity optionally provided through client certificate
- Integrity of exchanged data verified through Message Authentication Codes.

Attacks

High-profile protocol, thus valuable target.

- Protocol-version downgrades (FREAK and Logjam).
- Compression-based (CRIME and BREACH).
- Padding oracle-based (POODLE and Lucky Thirteen).
- ...



Our Contribution

- Differential testing of implementations.
- Combinatorial testing of X.509 certificate parsers:
  - All libraries should parse certificates the same way
  - Endpoint equivalence undecidable
  - Different behavior between implementations
    ⇒ possible vulnerabilities
- Combinatorial (sequence) testing:
  - Focus on handshake or entire TLS session
  - Hierarchical Input Parameter Models
  - Weighted t-way sequences
  - AI-based planning support in test generation

Target Implementations

- OpenSSL
- GnuTLS
- NSS
- ...

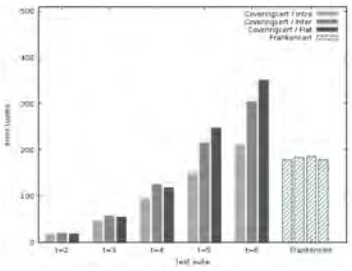
Contributions

CoveringCerts

Combinatorial generation of X.509 certificates and differential testing of parsers.

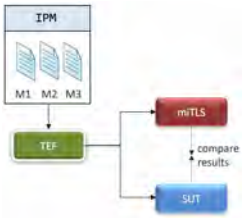
- Modeling of certificate contents.
- Generation of concrete certificates.
- Differential testing of implementations

⇒ More detailed and efficient results than previous approaches.



Hierarchical Input Parameter Models

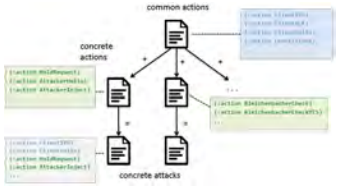
- Naive/flat approach:
  One model for all attributes of all messages in TLS handshake.
- Hierarchical approach:
  Intra-message model for each message,
  Inter-message model to combine results
  ⇒ Enables higher-strength testing.
- Comparison with miTLS, a verified reference implementation of TLS.



Sequence Testing

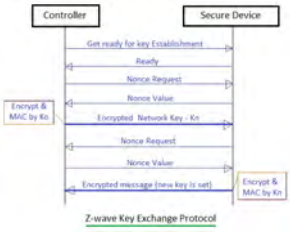
Modify attributes and order of TLS messages in handshake.

- Handshake testing as a sequence testing problem.
- Differential testing of implemented TLS state machines.
- AI-based planning to generate attack sequences.
- Weighted t-way Sequences:
  - Assign weights, derived from occurrences in bug reports, to events (TLS messages)
  - Event selection for candidate sequences based on integer partitions



Future focus: Internet of Things

- Rapid adoption:
  - Home automation / Smart Home
  - Medical assistance
  - Infrastructure management
- Resource constrained devices.
- Emerging protocols:
  - Z-Wave
  - NFC/RFID
  - Bluetooth Low Energy Mesh



Multi-faceted Attack Surface

- Attacks on web interfaces.
- Commonly backed by REST services.
- Focus on usability.
- Weakened cryptography.
- Increased privacy risk.

Rapidly changing technology
⇒Automated testing required.

- K. Kleine and D. E. Simos, "Coveringcerts: combinatorial methods for x. 509 certificate testing," in *2017 IEEE International conference on software testing, verification and validation (ICST)*. IEEE, 2017, pp. 69–79.
- D. E. Simos, J. Bozic, F. Duan, B. Garn, K. Kleine, Y. Lei, and F. Wotawa, "Testing tls using combinatorial methods and execution framework," in *IFIP International Conference on Testing Software and Systems*. Springer, 2017, pp. 162–177.
- B. Garn, D. E. Simos, F. Duan, Y. Lei, J. Bozic, and F. Wotawa, "Weighted combinatorial sequence testing for the tls protocol," in *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. IEEE, 2019, pp. 46–51.
- D. E. Simos, J. Bozic, B. Garn, M. Leithner, F. Duan, K. Kleine, Y. Lei, and F. Wotawa, "Testing tls using planning-based combinatorial methods and execution framework," *Software quality journal*, vol. 27, no. 2, pp. 703–729, 2019.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
Land Niederösterreich
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

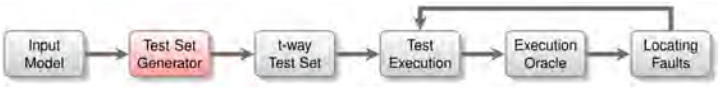
Covering Arrays Generation

Dimitris E. Simos, Michael Wagner, Ilias Kotsireas, Rick Kuhn, Raghu Kacker

Combinatorial Test Set Generation

Combinatorial Testing

- Combinatorial Testing allows for efficient testing of large systems while maintaining certain coverage guarantees.
- In a combinatorial test set, every t-way interaction appears in at least λ tests, where t is called the strength and λ the index.
- In practice, greedy algorithms have proven the most versatile approach and are therefore used in many combinatorial test generators.



Requirements for Combinatorial Test Generation Tools:

- Fast generation
- Small number of tests
- Easy to use
- Many different features

FIPOG and Tie-Breaker Evaluation

Various algorithmic and implementation-level improvements to the well-known In-Parameter-Order family of algorithms, including

- Simultaneous coverage gain computation.
- Skipping of fully covered column configurations.
- Partitioning of suitability checks.
- Compile-time strength.

Result: Vastly improved generation times.

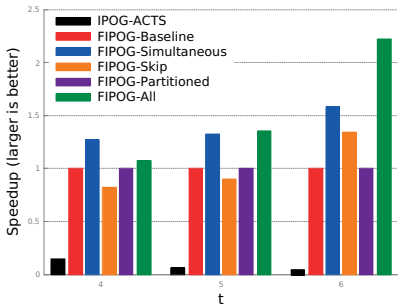


Figure 1: Speedup of FIPOG compared to the ACTS implementation of IPOG

- Further, different tie-breaking strategies were evaluated.

The CAgen Web GUI

- In the Input Parameter Model tab, the model can be edited.

CAgen: A tool for Fast t-Way Test Set Generation

- t-way test set generation up to strength $t = 8$.
- Implements the FIPOG, FIPOG-F and FIPOG-F2 algorithms.
- Support for constraints.
- Generation of test sets of higher index.
- Various export and import options.
- Compatible with other generation tools.
- Freely available as Web GUI and CLI at `https://matris.sba-research.org/tools/cagen`.

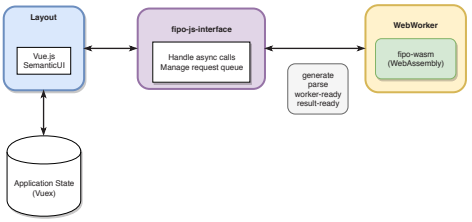


Figure 2: The architecture behind CAgen

Future Work and HPC

Optimization Algorithms

- Devise new metaheuristic algorithms, mathematical constructions and post-optimization methods for CA generation.
- Combine our experience in the field to devise efficient hybrid heuristics.
  - Enhance greedy algorithms using metaheuristics.
  - Combine mathematical constructions and reductions with other generation techniques.
  - Use Artificial Intelligence to enhance heuristic methods.
  - Develop a hyper-heuristic framework.

High Performance Computing

- Develop scalable parallel algorithms.
- Use super computing for constructing combinatorial test sets.

Charles J. Colbourn, Ryan E. Dougherty, Kristoffer Kleine, Dimitris E. Simos, and Michael Wagner. Algorithmic methods for covering arrays of higher index. Technical report, 2020.
Kristoffer Kleine, Ilias Kotsireas, and Dimitris E. Simos. Evaluation of tie-breaking and parameter ordering for the ipo family of algorithms used in covering array generation. In Costas Iliopoulos, Hon Wai Leong, and Wing-Kin Sung, editors, *Combinatorial Algorithms*. Springer International Publishing, 2018.
Kristoffer Kleine and Dimitris E. Simos. An efficient design and implementation of the in-parameter-order algorithm. *Mathematics in Computer Science*, 12(1):51–67, 2018.
Michael Wagner, Kristoffer Kleine, Dimiris E. Simos, Richard Kuhn, and Raghu Kacker. Cagen : A fast combinatorial test generation tool with support for constraints and higher-index arrays. In *To be published in Proceedings of IWCT 2020*. Springer International Publishing, 2020.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Covering Arrays, Set Covers, Algorithms and their Complexity

Dimitris E. Simos, Ludwig Kampel, Manuel Leithner, Bernhard Garn

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Generation of Covering Arrays for Abstract Combinatorial Test Suites

Covering Arrays for Combinatorial Testing

- Covering Arrays (CAs) provide the theoretical means for Combinatorial Testing (CT)
- Columns of a CA map to the parameters of a system under test.
- Rows of a CA encode the individual test cases.
- Their combinatorial properties guarantee that derived test sets **cover** all t**-way interactions**.
- To apply CT to arbitrary SUTs, we need to be able to generate arbitrary CAs.

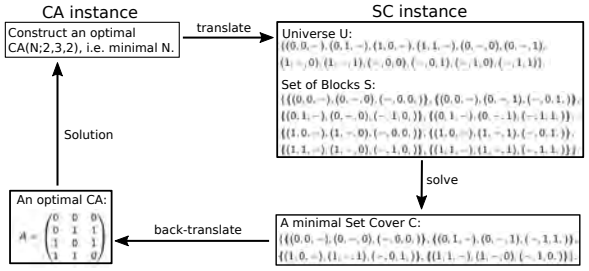
The Covering Array Generation Problem

- Given a **strength** t, a number of columns k and the respective columns' alphabet sizes $v_1, \ldots, v_k$.
- Construct a (mixed) covering array $\mathbf{MCA}(N; t, k, (v_1, \ldots, v_k))$ minimizing the number of rows N.
- Exact and direct constructions of CAs exist only for some corner cases.
- **For general applications we need heuristic algorithms for arbitrary CA generation.**

Covering Arrays via Set Covers

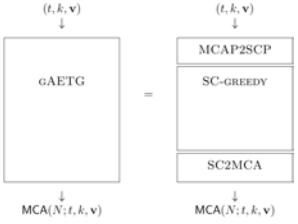
Optimal Covering Arrays as Minimal Set Covers

- The Set Cover Problem is a well studied problem in theoretical CS.
- For a given universe U and a set of blocks $\mathcal{S}$, i.e. subsets of U, we want to find a minimal subset of $\mathcal{S}$ that covers U.
- The CA generation problem can be interpreted as a Set Cover problem:
  - $U := \mathbb{T}_t$ the set of all t-way interactions
  - $\mathcal{S} := \prod_{i=1}^{k} [v_i]$ set of potential rows
  - Then a **minimal set cover** represents an **optimal CA**



Algorithms for Covering Arrays via Set Covers

- This connection allows to apply Set Cover (SC) Algorithms for CA generation.
- Some existing algorithms for CA generation can be identified as classical SC algorithms applied to CA instances.



- Allowing to import approximations and bounds for CAs:

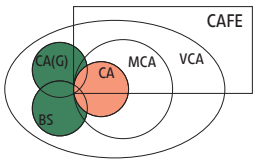
$$N \leq \mathit{MCAN}(N; t, k, \mathbf{v}) \cdot \log \binom{k}{t}$$

- This connection can be generalized to **weighted budgeted** instances pertaining to **weighted budgeted CA construction.**

Covering Arrays and Computational Complexity

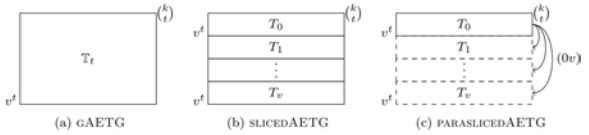
- Formulation of CA-related problems as formal complexity problems.
- Establish connections between these problems:
  - For arbitrary but fixed t and v, it holds that
    - (i) $\mathit{decSizeOMCA}_{t,v} \leq_p^T \mathit{detSizeOMCA}_{t,v} \leq_p^T \mathit{genOMCA}_{t,v}$.
    - (ii) $\mathit{decSizeOMCA}_{t,v} \equiv_p^T \mathit{detSizeOMCA}_{t,v}$.
- Analyse state of the art of complexity problems related to CAs.
- Correction of statements and clarification of misinterpretation.
- The computational complexity of the Covering Array generation problem remains unknown.

Classes of Covering Arrays	Decide Existence	Decide Size	Determine Size	Generation
optimal $CA_{2,2}$	P	P	P	P
optimal $CA_{t,v}$	P	NP	???	???
optimal $MCA_{t,v}$	P	NP	???	???
optimal $BS_{t,v}$	P	NP-complete	NP-hard	NP-hard
optimal $VCA_{t,2}$	P	NP-complete	NP-hard	NP-hard
optimal $VCA_{t,v}$	P	NP	???	???
optimal VCA_{*}	P	NP-hard	NP-hard	NP-hard
optimal CA(G)	P	NP-complete	NP-hard	NP-hard
CAFE	NP-complete	NP-hard	NP-hard	NP-hard



slicedAETG: A specialized Algorithm for CA construction

- The CA generation problem has more inherent structure compared to the general Set Cover problem.
- This can be exploited in order to devise more efficient algorithms.
- The slicedAETG algorithm is a specialization of a general greedy algorithm that is tailored to suite the CA generation problem.



Schematics how different algorithms process the set of all t-way interactions $\mathbb{T}_t$.

	$\geq$ # Rows	Runtime	Memory for $\overline{T}$
gAETG	$v^t \ln(v^t\binom{k}{t}) + 1$	$O\left(v^{k+t}\binom{k}{t}\ln(v^t\binom{k}{t})\right)$	$\Theta(v^t\binom{k}{t})$
slicedAETG	$v^{t+1}\ln(v^{t-1}\binom{k}{t}) + v$	$O\left(v^{k+t}\binom{k}{t}\ln(v^{t-1}\binom{k}{t})\right)$	$\Theta(v^{t-1}\binom{k}{t})$
paralslicedAETG	$v^{t+1}\ln(v^{t-1}\binom{k}{t}) + v$	$O\left(v^{k+t-1}\binom{k}{t}\ln(v^{t-1}\binom{k}{t})\right)$	$\Theta(v^{t-1}\binom{k}{t})$

Bounds on number of rows of output CAs, runtime and memory usage.

Ludwig Kampel, Bernhard Garn, and Dimitris E. Simos. Covering arrays via set covers. *Electronic Notes in Discrete Mathematics*, 65:11 – 16, 2018. 7th International Conference on Algebraic Informatics (CAI 2017): Design Theory Track.
Ludwig Kampel, Manuel Leithner, Bernhard Garn, and Dimitris E. Simos. Problems and algorithms for covering arrays via set covers. *Theoretical Computer Science*, 800:90 – 106, 2019. Special issue on Refereed papers from the CAI 2017 conference.
Ludwig Kampel, Manuel Leithner, and Dimitris E Simos. Sliced aetg: a memory-efficient variant of the aetg covering array generation algorithm. *Optimization Letters*, 2019.
Ludwig Kampel and Dimitris E. Simos. A survey on the state of the art of complexity problems for covering arrays. *Theoretical Computer Science*, 800:107 – 124, 2019. Special issue on Refereed papers from the CAI 2017 conference.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

(K)ERIS: A Novel Approach for API Security Testing, Applied to the System Call Interface of the Linux kernel

Dimitris E. Simos, Bernhard Garn

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Combinatorial Designs meet Software Testing and Information Security

Motivation

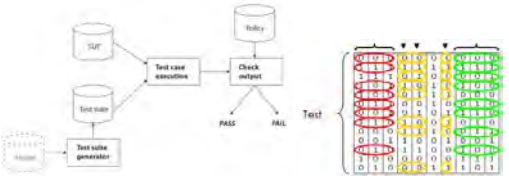
- We cannot test everything.
- Exhaustive search of problem space increases time needed exponentially.
- Automated detection of security vulnerabilities.

Combinatorial Security Testing (CST)

- Parameters and values provide abstract models of attacks.
- Generated test sets provide 100% coverage of t-way parameter value combinations.
- Automated test set generation, execution and evaluation via dedicated test oracle.

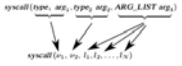
Technical Challenges

- Generation of minimal t-way test sets is a hard combinatorial optimization problem.
- Modelling of parameters, values and constraints is domain-specific.
- Deploy CST to all application layers of information security.



Combinatorial API Testing

- **Focus:** Test APIs function calls of software / libraries.
- **Modeling:** Combinatorial models:
  - IPM via equivalence- and category partitioning
  - IPM via novel flattening methodology
- **ERIS:** Highly configurable testing framework encompassing CT, execution environment, logging and database infrastructure.



Abstr. Parameter	Parameter values
ARG_CPU	1, 2, 3, 4, ..., 8
ARG_MODE_T	1, 2, 3, 4, …, 4095, 4096
ARG_PID	-3, -1, $pid_cron, $pid_w3m, 999999999
ARG_ADDRESS	null, $kernel_address, $page_zeros, $page_0xff, $page_allocs, …
ARG_FD	fd_1, fd_2, fd_3, ..., fd_{15}
ARG_PATHNAME	$pathname_1$, $pathname_2$, $pathname_3$, ..., $pathname_{15}$

ERIS: Combinatorial Kernel Testing

- **Focus:** Reliability and quality assurance of kernel software.
- **Motivation:** Kernel is the central authority to ensure security.
- **SUTs:** System calls of every git-commit of any (variant of) Linux.
- **Evaluation**: Various kernel crashes for RCs and distribution kernels.

Large-Scale Kernel Testing

Case Study

- Total of 3082 systems-under-test:
  - 23 different system calls
  - 134 kernel versions
- Kernel versions tested in the range of v4.0 up to v4.6:
  - The final releases
  - All release candidates
  - A selection of stable releases
- 102h execution time.

Evaluation via Differential Testing

- Compare number of accepted vs rejected system calls between versions.
- Mostly stable behaviour between versions.
- Largest deviations in the **settimeofday** system call:

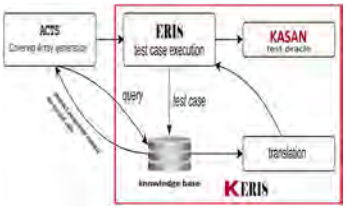
Count kernel versions	# of test cases	# of accepted	# of rejected
72	100	0	100
43	100	45	55
15	100	30	70
1	100	34	66

Evaluation via Kernel Address Sanitizer (KASAN)

- Test oracle uses internal dynamic memory error detector of Linux.
- Fine-tuned combinatorial model of a network configuration setup.
- Demonstrated reproducibility of vulnerability in the **sendto** system call.

Automated Test Execution Framework

- **Ease of use:** Only high-level parameters needed, everything else handled by the system.
- **Test-runs:** Each invocation runs in a dedicated virtual machine.
- **Logging:** Extensive information is captured:
  - Adjustable to user demands / needs
- **Database:** Allows sophisticated post-processing queries.



Vision

- **Goal:** Extend approach.
- **Modeling**: Optimization and automation of testing.
- **Automated** t-way testing and translation layers.
- Testing of **security patches** to ensure attack-free environments.
- **Continuous integration** tests of kernel versions.
- Web **monitoring platform**.

B. Garn and D.E. Simos. Eris: A Tool for Combinatorial Testing of the Linux System Call Interface. In *Software Testing, Verification and Validation Workshops (ICSTW), 2014 IEEE Seventh International Conference on*, pages 58–67, March 2014.
Bernhard Garn, Fabian Würfl, and Dimitris E. Simos. Keris: A ct tool of the linux kernel with dynamic memory analysis capabilities. In Ofer Strichman and Rachel Tzoref-Brill, editors, *Hardware and Software: Verification and Testing*, pages 225–228, Cham, 2017. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Large-scale Combinatorial Testing with Adobe

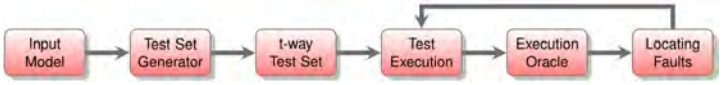
Dimitris E. Simos, Manuel Leithner, Ludwig Kampel, Darryl Jarman, Riley Smith, Jared Bellows, Rick Kuhn, Raghu Kacker

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Combinatorial Testing

Combinatorial Testing

- Combinatorial Testing allows for efficient testing of large systems while maintaining certain coverage guarantees.
- In a combinatorial test set, every t-way interaction appears in at least one tests, where t is called the strength of the test set.
- Studies have shown that the majority of faults can be detected with combinatorial test sets of strength $t \leq 6$.



Challenges for Large-scale Combinatorial Testing:

- Modelling of the system under test.
- Construction of reasonably small combinatorial test sets for systems with a large number of parameters is a difficult optimization problem.
- Sufficient Input Bandwidth necessary.

Generation of Large Combinatorial Test Sets

Greedy test set generation with CAgen:

- Generate test sets with a small number of rows with the FIPOG algorithm
- Significantly faster generation than other state of the art tools

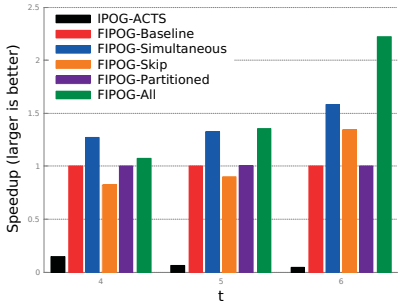


Figure 1: Speedup of CAgen compared to the state-of-the-art tool ACTS

- However: For the generation of combinatorial test sets for large systems, using greedy algorithms is no longer feasible.

Roux-type construction:

- Covering Arrays (CAs) are the mathematical structure underlying combinatorial test sets.
- Doubles the number of columns of a given CA.
- Concatenates permutations of CAs of lower strength to ensure coverage of all t-way interactions.

$$E_1 = \begin{pmatrix} A \\ A \end{pmatrix},$$

$$E_2 = \begin{pmatrix} B & B & \cdots & B \\ B^{\pi^1} & B^{\pi^2} & \cdots & B^{\pi^{r-1}} \end{pmatrix},$$

$$E_3 = \begin{pmatrix} C & C & \cdots & C & C^{f_1} & C^{f_2} & \cdots & C^{f_{N_1}} \\ C^{f_1} & C^{f_2} & \cdots & C^{f_{N_1}} & C & C & \cdots & C \end{pmatrix},$$

$$E_4 = \begin{pmatrix} C & C & \cdots & C & C^{g_{1,2}} & C^{g_{1,3}} & \cdots & C^{g_{t,r-1}} \\ C^{g_{1,2}} & C^{g_{1,3}} & \cdots & C^{g_{t,r-1}} & C & C & \cdots & C \end{pmatrix},$$

$$E_5 = \begin{pmatrix} C & C & \cdots & C & C^{g_{1,2}} & C^{g_{1,3}} & \cdots & C^{g_{t,r-1}} \\ C^{g_{1,2}} & C^{g_{1,3}} & \cdots & C^{g_{t,r-1}} & C & C & \cdots & C \end{pmatrix},$$

$$E_6 = \begin{pmatrix} C & C & \cdots & C & C^{h_{1,2}} & C^{h_{1,3}} & \cdots & C^{h_{t,r-1}} \\ C^{h_{1,2}} & C^{h_{1,3}} & \cdots & C^{h_{t,r-1}} & C & C & \cdots & C \end{pmatrix}.$$

Two-stage approach:

- Combines greedy methods with a mathematical doubling construction.
- CAgen generates seed arrays with as many columns as possible.
- Roux-type construction extends the seed arrays to the desired combinatorial test sets.

Large-scale Combinatorial Testing at Adobe

- Collaboration with Adobe.
- Application of largest combinatorial test sets documented in research.
- We generated test sets for models with more than 2000 parameters and 10 values:

N	t	k	v
6337	4	2127	3
107514	5	2127	3
87669	5	2127	3
322	2	2127	7
7439	3	2127	7
688	2	2127	10
23422	3	2127	10

Adobe

Adobe Analytics

Testing Results

- New faults found in each subject system.

Table I. REZZER FAULTS DETECTED BY COMBINATORIAL TESTING

Fault Descriptions, Causes, and Resolutions		
Description	*Cause*	*Resolution*
Flag-type fields throw error	Undocumented value constraint	Update input space model
Event-type fields throw error	Undocumented format constraint	Update input space model
Parser throws error (CDS)	Undocumented value constraint	Update input space model
Parser throws error (JSON)	Undocumented format constraint	Add input validation
Invalid date fields interaction	Undocumented value constraint	Update input space model

Darryl Jarman, Riley Smith, Gabe Gosney, Ludwig Kampel, Manuel Leithner, Dimitris E. Simos, Raghu Kacker, and Rick Kuhn. Applying combinatorial testing to large-scale data processing at adobe. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 190–193, 2019.
Kristoffer Kleine and Dimitris E. Simos. An efficient design and implementation of the in-parameter-order algorithm. *Mathematics in Computer Science*, 12(1):51–67, 2018.
Riley Smith, Darryl Jarman, Jared Bellows, Richard Kuhn, Raghu Kacker, and Dimitris Simos. Measuring combinatorial coverage at adobe. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 194–197, 2019.
Michael Wagner, Kristoffer Kleine, Dimitris E. Simos, Richard Kuhn, and Raghu Kacker. Cagen : A fast combinatorial test generation tool with support for constraints and higher-index arrays. In *To be published in Proceedings of IWCT 2020*. Springer International Publishing, 2020.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Quantum-Inspired Algorithms for Covering Arrays

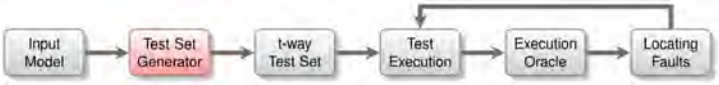
Dimitris E. Simos, Michael Wagner, Ludwig Kampel

COMET Competence Centers for Excellent Technologies www.ffg.at/comet
SBA Research
MATRIS

Covering Array Optimization

Covering Arrays

- Covering Arrays (CAs) are combinatorial structures used in Combinatorial Testing.
- They guarantee that every t-way combination appears in at least one row (test).
- A uniform, binary Covering Array is denoted as $\mathbf{CA}(N;t,k)$, where N is the number of rows, t the strength and k the number of columns.
- CAs with the smallest number of rows possible are called optimal CAs.



The Covering Array Generation Problem

- Generating optimal CAs is tightly coupled to hard combinatorial optimization problems.
- Commonly used generation methods include greedy algorithms, mathematical constructions and metaheuristic approaches.
- We investigated how quantum-inspired methods can help in generating near-optimal CAs.

Quantum-Inspired Evolutionary Algorithms

- First quantum-inspired evolutionary algorithm for CA generation.
- We introduced and evaluated new Mutation and Rotation types.
- We were able to generate various optimal binary CAs for strengths $t = 2, 3, 4$.

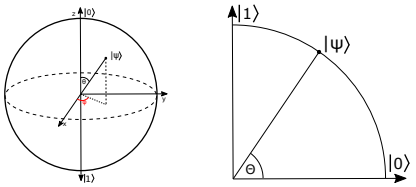


Figure 1: Qubit representation and the reduced version for real-valued amplitudes

Algorithm 1 QEAforCA(t, k, N)

Require: Rotation, Mutation, Termination
1: Create $Q(n)$ representing the $N \times k$ array
2: Create candidate solution $C(n)$ by observing $Q(n)$
3: Evaluate $C(n)$ based on the number of covered t-way interactions
4: $B(n) \leftarrow C(n)$
5: **while (not** Termination(B(n), t) **do**
6: $n \leftarrow n + 1$
7: Create $C(n)$ by observing $Q(n-1)$
8: **if** Evaluate $C(n)$ **then**
9: $B(n) \leftarrow C(n)$
10: **end if**
11: **for all** Qubits q_{ij} in $Q(n)$ **do**
12: $\alpha_{ij} \leftarrow$ Mutation(b_{ij}) ▷ Stops individual qubits from converging prematurely.
13: $q_{ij} \leftarrow$ Rotation(q_{ij}, α_{ij}) ▷ Updates the states of the Qubits to guide the search towards $B(n)$.
14: **end for**
15: **end while**
16: **return** $B(n)$

Example of the QEA Cycle

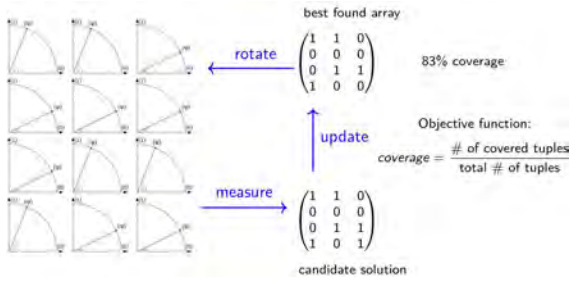


Figure 2: The workflow of the QEA cycle for the instance CA$(4; 2, 3)$.

Covering Array Generation using IPO-Q

- Combines QEA with the In-Parameter-Order strategy:

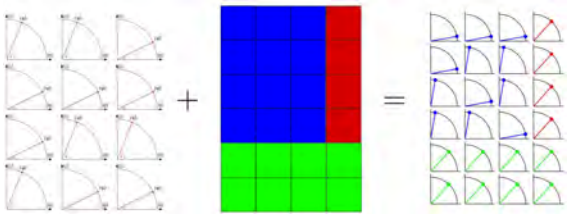


Figure 3: IPO-Q: Combining QEA with IPO

- Expands array using vertical and horizontal extension steps:
  - The blue Qubits in Figure 3 represent the CA from the previous extension step, their state biased towards their old value
  - A new column is added (red Qubits) and QEA attempts to find a CA with the newly added column
  - If QEA fails to generate a CA, additional rows are added (green Qubits)

IPO-Q Evaluation and Future Work

- Guaranteed CA upon termination
- Improved on other IPO variants by reducing the number of rows for certain binary CA instances:

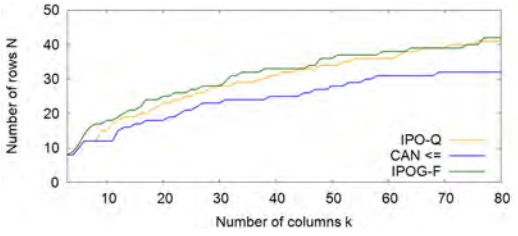


Figure 4: Comparsion between the number of rows generated by IPO-Q and IPOG-F for binary CAs of strength $t = 3$

Future Work:

- Generalize our QEA and IPO-Q for higher alphabets.
- Use Quantum Computing to solve Covering Array problems.

Michael Wagner, Ludwig Kampel, and Dimitris E. Simos. Quantum-inspired evolutionary algorithms for covering arrays of arbitrary strength. In Ilias Kotsireas, Panos Pardalos, Konstantinos E. Parsopoulos, Dimitris Souravlias, and Arsenis Tsokas, editors, *Analysis of Experimental Algorithms*, pages 300–316, Cham, 2019. Springer Publishing.
Michael Wagner, Ludwig Kampel, and Dimitris E. Simos. Ipo-q: A quantum-inspired approach to the ipo strategy used in ca generation. In Daniel Slamanig, Elias Tsigaridas, and Zafeirakis Zafeirakopoulos, editors, *Mathematical Aspects of Computer and Information Sciences*, pages 313–323, Cham, 2020. Springer International Publishing.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Beyond Research

ISS – Consulting & Trainings

SIC – Strategic Innovation & Communication

Information Security Services
Research and Consulting under One Roof

Stefan Jakoubi

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

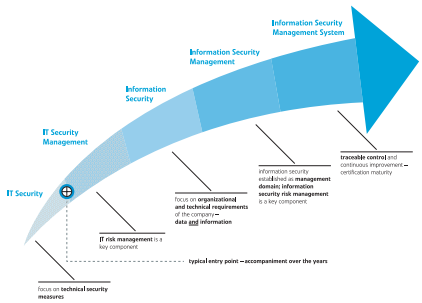
From Quality Assurance to Structured Support for an ISO27001 Certification

Information Security Services is the consulting department of SBA Research, which has been appreciated for years as a reliable partner of ministries, public authorities, large companies, and SMEs.

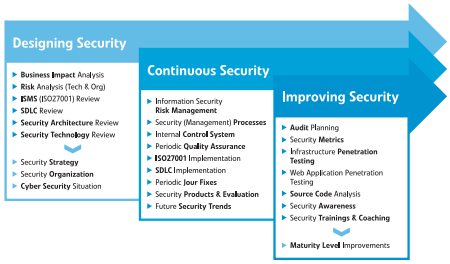
Our organizational and technical security analyses and consulting services contribute significantly to increase the security level of our corporate partners. This **dual approach of scientific research and practice-oriented implementation** enables a unique range of services: from research collaborations to penetration tests to covering security aspects of future key areas such as AI, IIoT/Industry 4.0, secure software development and security in digitalization. This is supplemented by a comprehensive portfolio of training courses.

Thereby, we pursue three core objectives:

1. **Long-term support** of company partners who are at the beginning of their journey to the desired security maturity level.
2. **Specialized analyses** and quality assurance for corporate partners as part of their continuous improvement process.
3. **Comprehensive knowledge transfer**, as employees are the most valuable asset and „first line of defense" of any company.



We accompany our partners on their **way to an information security management system ...**



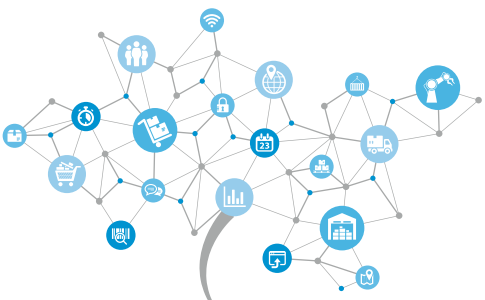
... and offer consulting and analyses in the **entire information security spectrum**.

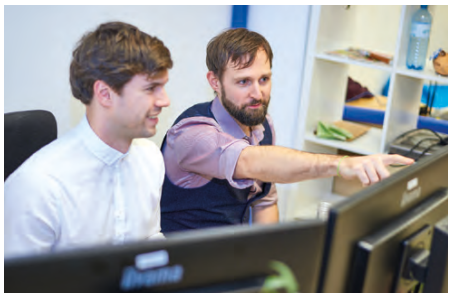
Every year, we support around 70 corporate partners, carry out approx. 150 projects and train hundreds of participants in our course and knowledge transfer programs. Our approx. 20 information security consultants accompany our clients with longstanding professional experience.

Understanding a Company is the Key for the Security Strategy

We analyze the (digital) ecosystem

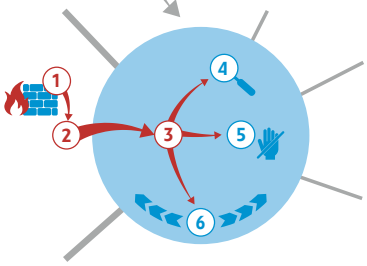
- In what an **environment** does the company operate?
- What is the **product?** Primary markets?
- What do the **employees** need?
- **Dependencies** on suppliers?
- **Competition?**
- Legal and regulatory **requirements?**
- Sector-specific **threat situations?**
- **Business impact** of an incident?





We simulate the attackers ...

1. **How secure is the „castle wall"?**
2. **Where are the holes in the wall?**
3. **Can an attacker spread out on the inside?**



... and strengthen the defenders!

4. **Can the attack be detected?**
5. **Can the risk of data loss be reduced?**
6. **How good are the security processes? Are the employees well trained?**

SBA attaches great importance to building bridges between top-level research and practical results and considers itself – thanks to its extensive national and international network – as a link between science and industry.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien Für die Stadt Wien
European Commission
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Advanced Trainings @ SBA Research

Effective cyber defense and reaction starts with expertise

Information Security Services

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

Advanced Trainings @ SBA Research

In cyber security, employees are the first line of defense. Due to a continuing shortage of specialists, companies must strongly engage in employee training programs.

SBA Research offers such specialized training courses. Our trainers are certified security experts and highly experienced. We provide in-house as well as on-site trainings which comply with state-of-the-art best practices and standards.

Over the last 10 years, we have trained **approx. 900 participants in about 150 publicly available trainings**, not counting our numerous customized on-site courses.

SBA's Trainings Portfolio – An Overview

Security Starts with Awareness	Deep Dive Trainings	Preparation Courses for InfoSec Certifications
Security Awareness (1.5–2h, 20 participants)	Windows Hacking (3 days, 12 participants)	CSSLP (5 days, 12 participants)
Cyber Security Essentials (3 days, 12 participants)	Web Application Security (3 days, 12 participants)	CISSP (5 days, 12 participants)
	Incident Response (3 days, 12 participants)	CISM (4 days, 12 participants)
	IoT Security Essentials (3 days, 12 participants)	CISA (4 days, 12 participants)

Cyber Security Essentials & Awareness

An introduction to cyber-, IT-, and information security

TOPICS
- Cyber security, current threats, and risk scenarios
- Basics of IT- and information security
- Legal and organizational framework
- Solutions and measures to deal with cyber risks
- Challenges and pitfalls in practice

TARGET GROUP
- People in management positions
- Project managers
- Demand managers
- Security „ambassadors" in all departments

KEY TAKE-AWAYS
- Experience a journey through the "cyber sec world"
- Gain a basic knowledge of cyber security
- Develop an awareness of essential threats, trends, standards, and security measures

IoT Security Essentials

The essential aspects when evaluating an IoT device and assessing its security status

TOPICS
- Firmware analysis & reverse engineering
- Step-by-step development of an exploit
- Bus systems, interfaces, prototyping of interfaces
- Solutions and measures to cope with attacks
- Challenges and pitfalls in practice
- Additionally apply learned techniques in in-depth, hands-on labs

TARGET GROUP
- Software developers
- System administrators without knowledge of secure development

KEY TAKE-AWAYS
- Learning the content based on a fictitious scenario
- Acquire sound fundamental knowledge of cyber security
- Develop an awareness of essential threats, trends, standards, and security measures

Windows Hacking

Overview of the most common and dangerous vulnerabilities in Windows networks

TOPICS
- Security vulnerabilities and their mitigation for Windows clients
- Vulnerabilities and related security measures for Windows servers
- Attack points and protective measures in the network and on mobile devices (laptop)
- Additionally apply learned techniques in in-depth, hands-on labs

TARGET GROUP
- Windows Administrators & Windows networkers
- People in IT management

KEY TAKE-AWAYS
- Major threats in Windows networks
- Techniques and tools for the detection and mitigation of vulnerabilities

Web Application Security

Avoid the most dangerous pitfalls in web application development

TOPICS
- Typical and dangerous web vulnerabilities (OWASP)
- Techniques and tools for the detecttion and mitigation of vulnerabilities
- Independent of specific programming languages
- Additionally apply learned techniques in in-depth, hands-on labs

TARGET GROUP
- Development teams of web applications
- Test teams of web applications

KEY TAKE-AWAYS
- How to detect and defend against typical web application vulnerabilities
- Learn how to develop secure web applications

Incident Response

Techniques and tools for the investigation of a cyber attack

TOPICS
- Simulation scenario based on a real attack
- Analysis of phishing e-mails
- Triangulation of attacks in the network based on log files
- Evaluation of live data
- Creation and evaluation of memory images and hard disk images
- Creating a timeline

TARGET GROUP
- Incident Response Team Members
- System administrators
- Persons responsible for information security with technical background

KEY TAKE-AWAYS
- Learning based on a realistic scenario
- Learn techniques and tools to detect, contain and mitigate a cyber attack

Preparation Courses for IT Security Certifications

Exam preparation courses in four areas:

- **Certified Information Systems Security Professional (CISSP)**
  Security & Risk Management | Asset Security | Security Engineering | Communication & Network Security | Software Development Security
- **Certified Secure Software Lifecycle Professional (CSSLP)**
  Concepts & Requirements | Design & Implementation/Coding | Testing | Software Acceptance
- **Certified Information Systems Auditor (CISA)**
  Auditing Process | Governance and Management of IT | Acquisition | Business Resilience | Protection of Assets
- **Certified Information Security Manager (CISM)**
  InfoSec Governance | Information Risk Management & Compliance | InfoSec Program Development & Management | Security Incident Management

TARGET GROUP
Depending on the certification, from all interested persons to experts in the respective field.

KEY TAKE-AWAYS
- Thorough exam preparation
- Discussion of practical experience & best practices

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien Für die Stadt Wien
European Commission
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

How to Do Research with SBA Research

Veronika Nowak

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

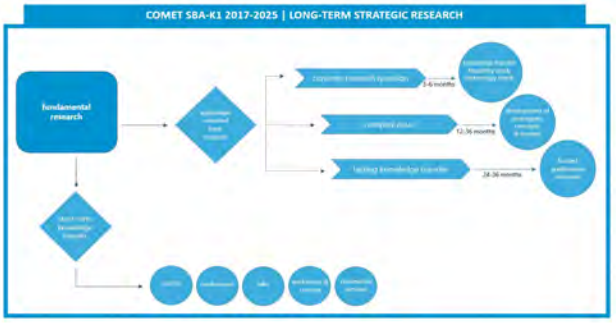
SBA Research

SBA Research provides top-level basic and application-oriented research on a national as well as international level. First and foremost, we are pursuing an **ambitious long-term strategic research** agenda, divided into four main areas. Based on this fundamental research, we further offer **application-oriented basic research**, executed in joint projects with our company partners.

COMET strategic research
long-term fundamental research

- breaking new scientific ground
- contributing to a secure society

Area 1: Networked Systems Security
Area 2: Software Security
Area 3: Privacy and Secure Societies
Area 4: Applied Discrete Mathematics for Information Security



*We strive to establish **trust in cyberspace (Area 1)** and **machine learning-based software systems (Area 2)**, considering **privacy and societal aspects (Area 3)** through innovative approaches like utilizing **discrete mathematics for information security challenges (Area 4)**.*

COMET Partnership

long-term cooperation

- immediate implementation of research projects
- completely flexible regarding the research topic

SBA Research has over 40, mostly long-standing COMET company partners of which well over a third are small and medium enterprises.

In joint research projects with the COMET Center SBA Research we are able to design new methodical analysis procedures that serve cybersecurity awareness and resilience development, and to test these prototypically with other partners.
Frank Christian Sprengel, Repuco Unternehmensberatung GmbH

Short-Term Transfer of Knowledge

meeting immediate and hands-on needs for expertise

- conferences, events, lectures, workshops and trainings
- in-house, or on-site at our company partners and clients

Example
bridging the gap between software development and information security **sec4dev Conference & Bootcamp**, organized annually by SBA Research

Complex Research Issues

development of prototypes, concepts and models

- joint research project, 12 to 36 months
- top level research to be further developed into marketable products and services

Example
creating target device tailored honeypots for the (Industrial) Internet of Things which are capable of convincing an adversary that a real device was breached | FFG-funded, 2-year research project "AutoHoney(I)IoT"

Funding Programs (FFG)
- BRIDGE 1: near-basic research, 2 to 3 years
- ICT of the Future: applied research, 3 years

Concrete Research Question

supporting innovative ideas through technology checks and feasibility studies

- medium-term measures (e.g., technology checks): 3-6 months
- FFG funding for SMEs

Example
keeping sensitive health data safe – the HeartBalance health measurement device

The cooperation with SBA enables our company to develop our products and services in a highly future-oriented way.
Peter Hauschild, HeartBalance Innovations GmbH

Funding Programs (FFG)
- Innovation Check with Deductible: purchase of additional research services with a maximum value of EUR 12,500; EUR 10,000 grant
- Small Project: max. EUR 90,000 funding, duration: 18 months

Sustainable Knowledge Transfer

converting current research results into know-how for the company partners

- long-term training measures, 24 to 48 months
- jointly developed by all project partners

Example
information security in IT/OT environments, considering technical and organizational aspects of production from the point of view of attackers and defenders | FFG-funded, 2-year specialized training program "InduSec" with 8 industry- and 4 scientific partners

Thanks to InduSec, I can now ask the right questions regarding IT and OT infrastructure when working with external service providers, and I'm aware of security issues when migrating processes.
feedback from a participating SME

Funding Programs (FFG)
- Qualification Networks: specialized courses, 1 to 2 years
- Tertiary Level Courses: tailor-made trainings, 2 to 4 years

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Associations & Networks

Knowledge Exchange | Scientific Outreach | Education & Training | Awareness

Daniela Friedl, Stephanie Jakoubi, Barbara Limbeck-Lilienau

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

Associations & Networks

SBA Research is continuously building networks and contributes to numerous platforms, for example IEEE SMC/CS Austria, ACM SIGSAC Vienna, IFIP WG 8.4, and (ISC)² Chapter Austria. Memberships with other platforms like CONCORDIA, COST Actions, ECSEL Austria, EC3 Europol, ECSO, or Industrie 4.0 facilitate a constant **knowledge exchange with experts from academia and industry**, both at home and abroad. Young academics – from bachelor to post-doc – are furthered and integrated in existing networks via **summer schools** (e.g., IPICS) and **exchange programs** (e.g., ERCIM).

These strong networks form the basis for various core activities – ranging from **scientific outreach & knowledge exchange** and organizing **academic security conferences & security events** to forming **strong consortia** between academia and industry to **education & training** – thus boosting **awareness** and assuming **social responsibility**.

National and International Networks (selected)

List of Associations and Networks

Association | Network (international)	Role	Focus
ACM Association for Computing Machinery	Member	Research | Scientific Outreach | Academic Conferences
Concordia H2020 Competence Network	Partner	Research | Industry | Standardization | Knowledge Exchange
European Cooperation in Science and Technology (COST) Actions (DigForASP)	Member	Research | Knowledge Exchange | Scientific Outreach
European Cyber Security Organisation (ECSO)	Member	Industry | Research | Standardization | Public-Private Partnership
European Research Consortium for Informatics and Mathematics (ERCIM)	Member	Research | Exchange Program | Knowledge Exchange | Scientific Outreach
Europol's Cybercrime Center Academic Advisory Network (EC3AAN)	Member	Academic Advisory Network | Knowledge Exchange
Information Systems Audit and Control Association (ISACA)	Member	Industry | Knowledge Exchange | Security Conferences & Events
IEEE – Institute of Electrical and Electronics Engineers	Member	Research | Scientific Outreach | Academic Conferences & Events
Institute of Combinatorics and its Applications (ICA)	Fellow	Research | Knowledge Exchange
International Information Systems Security Certification Consortium – (ISC)²	Member	Industry | Knowledge Exchange | Certification
Intensive Program on Information and Communication Systems Security (IPICS)	Co-Founder	Research | Summer School | Knowledge Exchange
International Federation for Information Processing – WG 8.4 E-Business Information Systems (ifip)	Chapter Founder	Research | Scientific Outreach | Academic Conferences & Events
National Institute of Standards and Technology (NIST)	Cooperation	Standards and Guidelines | Joint Research Projects
Research Data Alliance (RDA) – Data Citation WG and DMP Common Standards WG	Co-Chair	Research | Standardization | Knowledge Exchange

Association | Network (national)	Role	Focus
ACM Special Interest Group on Security, Audit and Control – Vienna Chapter (ACM SIGSAC)	Chapter Founder	Research | Scientific Outreach | Academic Conferences
Austrian Computer Society (OCG)	Member | Chapter Co-Chair	Research | Industry | Knowledge Exchange | Academic Security Events
Austrian Standards – ON_AG 00188 Blockchain	Member	Standardization | Industry | Research | Knowledge Exchange
DCNA – Disaster Competence Network Austria	Member	Research | Industry | Knowledge Exchange | Education & Training
Decision Support for Health Policy and Planning (DEXHELPP)	Member	Research | Industry | Knowledge Exchange
DigitalCity.Wien	Board Member	Industry | Research | Knowledge Exchange | Awareness | Education
Electronic Components and Systems for European Leadership (ECSEL Austria)	Member	Industry | Research | Knowledge Exchange
IEEE Systems, Man, and Cybernetics Society (IEEE SMC/CS Austria)	Chapter Founder	Research | Scientific Outreach | Academic Conferences & Events
International Information Systems Security Certification Consortium – Austria Chapter (ISC)²	Chapter Founder	Industry | Knowledge Exchange | Security Conferences & Events
ÖFG – Digitale Transformation: Safety, Security & Privacy	Chapter Founder	Research | Scientific Outreach | Knowledge Exchange
Research Data Alliance Austria (RDA-AT)	Co-Chair	Research | Standardization | Knowledge Exchange
Verein Industrie 4.0	Member	Industry | Research | Knowledge Exchange

Concordia

Cyber Security Competence for Research and Innovation

Partner | European H2020 Network | Research | Industry | Knowledge Exchange | Awareness

- One of four H2020 projects that pilot an EU Cybersecurity Competence Network
- 50+ partners including prominent industry, academia, SMEs, and public bodies
- 20 countries (16 EU member states, 3 Horizon 2020 associated countries, and UK)
- 16m EU funding for 4 years, 7m additional funding from national authorities and industry
- Aim: retaining and developing the technological and industrial cybersecurity capacities necessary to advance Europe's Digital Single Market and to strengthen Europe's overall cybersecurity

DigitalCity.Wien

DigitalCity.Vienna

Member of the Board | Regional | Knowledge Exchange | Education | Awareness

- Cooperation on topics of information and communication technology which are of essential importance for the Smart City Vienna initiative
- Positioning and visibility of the topic „ICT and digital competence" in all training areas and institutes
- Strengthening Vienna as a digital ICT metropolis and innovative pioneer in the international ICT context
- Knowledge exchange with experts from academia and industry
- Digital Days: 4,500+ visitors, 120+ industry partners, 60+ Digi-Street-Stations, 70+ speakers & panelists (2019)

(ISC)² Austria Chapter

International Information Systems Security Certification Consortium

Chapter Founder | National | Industry | Knowledge Exchange | Security Conferences & Events

- Promotion of a network of information security experts with (ISC)² certifications in Austria
- Holding regular chapter events for exchanging/presenting new developments in the information security field and to share know-how with like-minded people
- Fostering communication between the international organization (ISC)², the Chapters in other countries and the local Chapters in Austria
- (ISC)² ISACA Conference: annual member event with more than 180 members of the Austrian Chapters, co-organized by (ISC)² Austria and ISACA Austria and hosted by SBA Research

The Research Data Alliance (RDA)

RDA DMP Common Standards WG | Research Data Alliance Austria (RDA-AT)

Chair & Coordinator | Internat. & Nat. | Research & Industry | Knowledge Exchange | Standardization

Photo: RDA-AT/Gerhard Mayer

- Building the social and technical bridges to enable the open sharing and re-use of data
- Community-driven, grass-roots, inclusive approach covering all data lifecycle stages; engaging data producers, users, and stewards; addressing data exchange, processing, and storage
- 10.000+ members from 144 countries
- RDA DMP Common Standards WG on machine-actionable data management plans (DMPs), 200+ members
- RDA-AT: linking Austrian data management initiatives and RDA Working and Interest Groups

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
Für die Stadt Wien
European Commission
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Security Research & Knowledge Transfer
Bringing IT- and Information Security to People

Thomas Konrad, Julia Pammer, Nicolas Petri, Yvonne Poul, Stefan Jakoubi, Stephanie Jakoubi

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

IT Security & Information Security from Research to Everyone

Besides scientific conferences and events, SBA takes pride in organizing and supporting a multitude of events for knowledge transfer and dissemination. Thus we are addressing the professional education of partner companies as well as strengthen and support the relationship with young researchers and students. Most of our partners and experts from the security community take advantage of our broad expertise and send their employees to our educational events.

- from security awareness to technical details
- building communities
- transferring research to best practice

Selected Initiatives & Activities

- **ARES Conference**, annual academic conference
- **sec4dev** Conference & Bootcamp: 4 days for secure software development
- **Trainings**, from awareness to preparation for certification exams
- Cyber Security Awareness Quiz
- Security awareness for students
- Expert talks at numerous external events
- Events hosted by SBA, 2010-2019: 89
- Building and fostering numerous **communities**
- Supporting the Bildungsinitiative of the City of Vienna
- Giving the stage to young researchers at the **Young Researchers Day at the IKT Sicherheitskonferenz**

Expert Talks & Round Tables

From overviews to deep dives

Sharing our expertise, passion and curiosity about security, considering various systems and technologies – for example:

- Security Awareness
- IoT, IIoT, Industry 4.0, AI
- Information Security
- Threat landscape

30+ talks/year | 6 research groups | 15+ different security topics

sbaPRIME

Trends – Insights – Guidance – Awareness

sbaPRIME is at the interface of research, best practice, know-how and business. We provide practical and application-oriented solutions, hands-on knowledge and best-practice methods in the field of information security.

security news | exclusive events | white papers | in-house meetings

Security Meetups by SBA Research

Secure Coding Community Austria

Our experts present and discuss common security pitfalls which they regularly encounter in live systems. They provide tips and best-practice methods to make applications more secure. Awareness for secure coding is growing in the community.

- Focus on secure software development
- Sharing our experience
- Security awareness
- Regular meetings
- Online and offline Meetups
- 500 members and growing

16 Meetups | 500 members

SBA Live Academy

30 Minutes for Security

A platform for hand-picked security expert talks:**Tailored** short talks covering **current security topics,** including Q&A with our experts. The talks have been recorded for a sustainable and long-term knowledge transfer.

25 events | 800 minutes of talks | 900 participants in total

Cyber Security Awareness Quiz

Security Awareness for Everyone

- Learning about cyber security
- Boost security knowledge
- Gamification
- Mobile and web application

SBA Research is a leading partner and designed the app to advance the security knowledge of Austria's citizens.

Impressions





SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

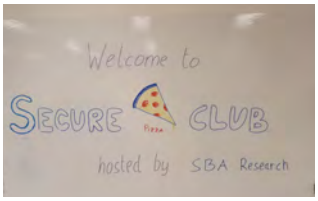
Students @ SBA Research (securepizza.club & CTFs)

Nicolas Petri, Wilfried Mayer, Georg Merzdovnik

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Motivation

- The club serves as **platform and get-together for students**.
- A place where students can watch security videos (DEF CON, Black Hat, etc.), talk about new security issues and **exchange ideas.**
- **Pizza & Drinks** are provided by SBA Research.
- The goal is to promote an improved and **strong relationship with students.**



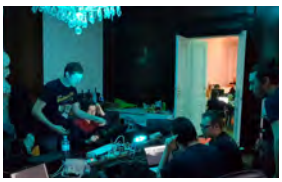
securepizza.club at SBA Research

- Up to 4 events a year
- 530+ slices of pizza
- 46 participants on-site
- Online events in the near future

CTF We_0wn_Y0u

We_0wn_Y0u is a CTF team of TU Wien students and SBA Research employees. We are participating in national and international online CTFs as often as possible and hold monthly meetings for our members.

- Founded: 2004
- Active users per month: 40
- Past CTF events (2017-2019): 67
- One of the top 3 CTF teams in Austria



Topics and Content

- Presentation of our research
- Covering a broad range of security concepts and topics, such as:

Image taken from WillC - Phreaking Elevators - DEF CON 27 Conference
https://www.youtube.com/watch?v=NoZ7ujJhb3k

- **SELECT Code Execution from Using SQlite**
  - simply querying a malicious SQLite database can lead to Remote Code Execution
  - created a rogue SQLite database that exploits the software used to open it
- **Adventures in Smart IoT Devices Penetration Testing**
  - insight into the world of smart IoT devices to see how connected tools' security holds up against a vaguely motivated attacker
- **Phreaking Elevators**
  - a comprehensive dive into current emergency phones with an in-depth look at the phones used in elevators

Outcome

- Interesting **discussions** about current information security topics
- **Interconnection** between research center and motivated students
- Showing **career opportunities** in information security research and -consulting
- Presenting **current research** areas of SBA Research
- Informing about different topics concerning **Bachelor, Master** and **PhD** theses
- Outlining the **bridges** between information security research and applied challenges
- Presenting **successful examples** of research- and commercial projects

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Conferences Hosted by SBA Research

Julia Pammer, Bettina Jaber, Yvonne Poul

COMET
Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Conference Hosting

SBA Research is regularly hosting international and national conferences. We provide profound expertise in scientifically curating as well as managing a variety of events.

Our focus is to further facilitate knowledge transfer by providing the platforms and creating an atmosphere that enable discussion and exchange. Additionally, SBA Research strives to bring top scientific conferences to Vienna, thus raising Austria's standing in the international research community.

Photo: Lecture Hall ARES 2019

ARES Conference

ARES Conference
International Conference on Availability, Reliability and Security

- Hosted annually since 2006, in 11 different countries so far, with 200 participants on average
- Highlights the various aspects of security – with special focus on the crucial linkage between availability, reliability, and security
- Up to 15 workshops & 8 EU Project Symposium workshops each year
- International IFIP Cross Domain Conference for Machine Learning & Knowledge Extraction (CD-MAKE 2020) co-located
- ARES is ranked as B-conference in CORE.
- Qualis (backed by Brazilian Ministry) ranked ARES and ESORICS as leading security conferences in Europe (A2).
- Acceptance rate for ARES full papers is ~ 20 %.

Photo: Group Picture ARES 2019

Photo: Mine visit in Berchtesgaden, Germany; ARES 2016

sec4dev Conference & Bootcamp for Software Developers

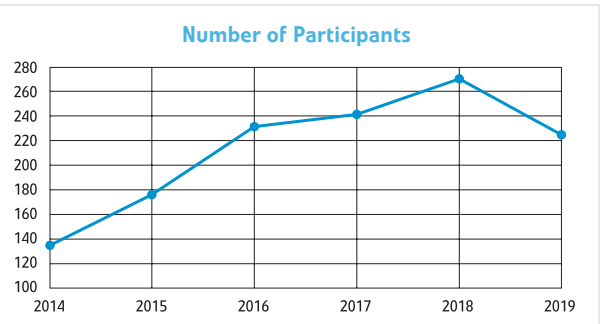
sec4dev

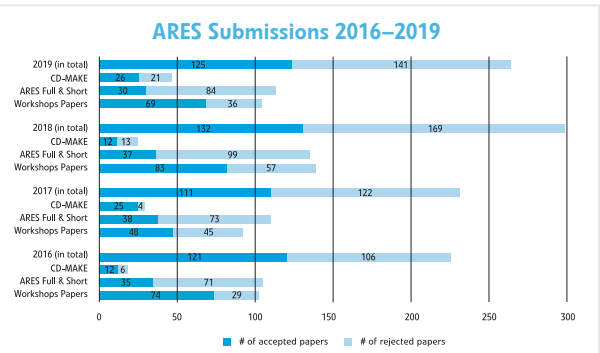
- Vienna-based security event which welcomes around 200 people each year involved in software development
- Launched in 2019 with the mission to make security a first-class citizen in software development by bringing together the best industry professionals and thus having a sustainable and positive impact on the software security landscape
- Focus on practical, applicable, hands-on, and security-related contents for people involved in software development
- Through sec4dev and regular Security Meetups, SBA Research creates and promotes a growing community which already includes over 600 software developers

Photo: Group Picture sec4dev 2020

Photo: sec4dev Team

ARES in numbers





Further Conferences Hosted by SBA Research

Including the annual ARES Conference and sec4dev Conference & Bootcamp, SBA Research hosted 32 international conferences between 2010 to 2019.

ACM Conference on Computer and Communications Security (CCS)

A particular highlight was the ACM Conference on Computer and Communications Security (CCS), one of the internationally most renowned flagship conferences in information security. In 2016, SBA Research hosted the ACM CCS at the Hofburg Palace in Vienna:

- more than 1,000 participants from over 40 countries
- 137 papers out of 831 submissions presented (16,5 % acceptance rate)
- 14 workshops, 7 tutorials, 3 industrial talks, 2 keynotes, 1 panel discussion

Photo: Keynote Speaker Ross Anderson

Photo: The "Coffeehouse" was very popular for exchanging ideas while enjoying the view of the Inner City of Vienna.

Photo: Panel Discussion CCS 2016

Photo: General Chair Edgar Weippl and Head of Organization Yvonne Poul (both SBA Research) in the popular photo booth

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
Für die Stadt Wien
European Commission
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

SBA's Appearance in Media

Mailyn Stolz

COMET
Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

Research Year 1 (27 articles in total)

"However, the software that is installed may itself have security vulnerabilities and allow new attacks"
– Edgar Weippl

"The darknet is not automatically illegal"
– Peter Kieseberg

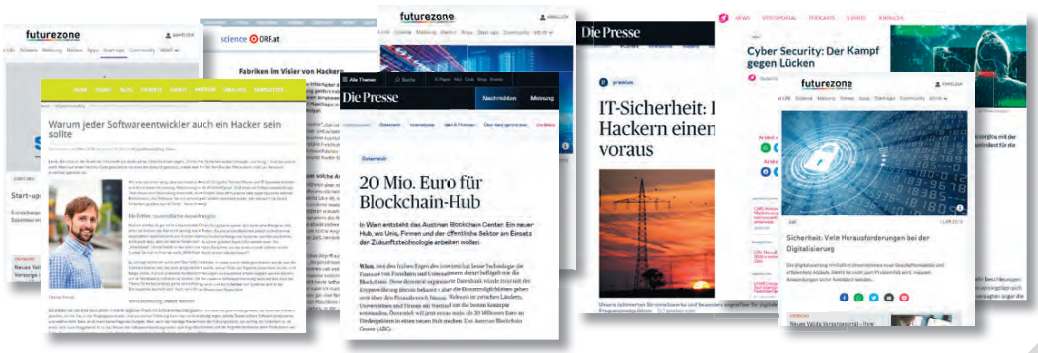
"Password managers are programs with which you can save your account data, i.e. access name and password"
– Edgar Weippl

Research Year 2 (17 articles in total)

"A Society needs trust to function"
– Edgar Weippl

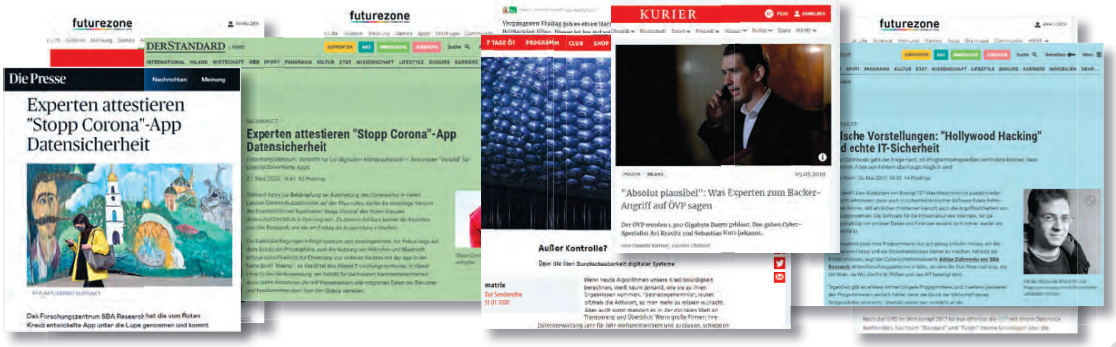
"Secure software is a matter of thinking. The later a fault is found, the more expensive it usually is to correct"
– Thomas Konrad

"I would think a mediocre computer scientist could do that"
– Johanna Ullrich



Research Year 3 (15 articles plus 25 articles only about the Corona-App Analysis)

"It is possible for attackers to recognize smartphones at specific locations over long periods of time and, in extreme cases, to create movement profiles. We have been promised that this problem will be fixed with a new version at the end of next week"
– Christian Kudera



Selected List of Media with SBA Coverage

der brutkasten · DERSTANDARD · Die Presse · ORF · heise online · futurezone · radio FM4 · Report · profil · WDR · NÖN · Heute · Kronen Zeitung · oe24 · Neues Volksblatt · SALZBURG 24 · OÖNachrichten · Tiroler Tageszeitung · VIENNA ONLINE · VORARLBERG ONLINE · KLEINE ZEITUNG · SÜDTIROL NEWS · DR

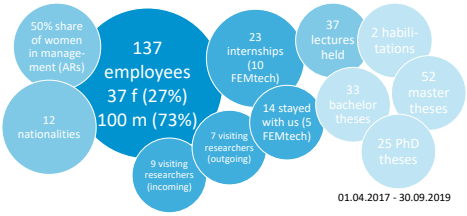
Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien
Für die Stadt Wien
European Commission
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Talents @ SBA Research

Belinda Reisinger-Nossek, Nicolas Petri

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

SBA Research in Numbers

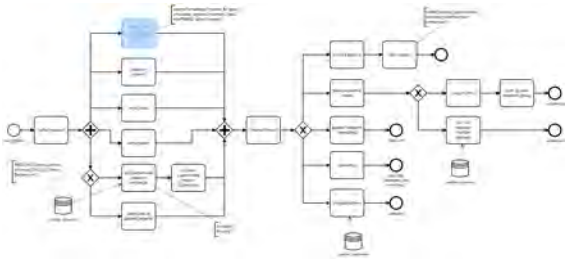


Employee Benefits

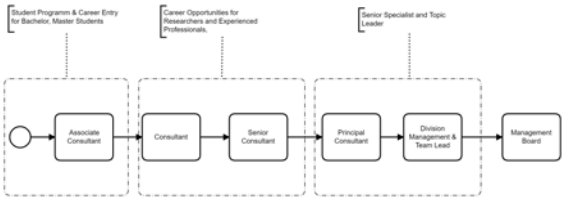
Home Office & Flextime (no all-in contract)	**Team- & Company Events Community Building**	**Manifold Career Opportunities**
Good Public Transport Connections	**Interdisciplinary Research Projects**	**SBA Fellow Program** Staying in touch with former SBA employees.
Coffee, Juice, Mineral Water, Fruit for Free	**Personnel Development (trainings, further education)**	

Career Opportunities

We offer our researchers manifold carrier prospects at SBA Research as well as with (inter-) national research institutions and COMET partners. Furthermore, we provide and support in-house and external opportunities of personnel development.



Our consultants are encouraged to become well-established allrounders with an information security specialization of their choosing:



Success Stories - FEMtech

We strongly encourage female students to pursue a career in information security research. Participating in the FEMtech program proved to be a valuable asset to our company, allowing us to hire several competent female researchers.
For instance, Katharina Krombholz and Johanna Ullrich started their scientific in research with a FFG-funded FEMtech internship and were subsequently offered employment at SBA Research.

In 2013, Katharina Krombholz spent a semester as a research intern at the National Institute of Informatics in Tokyo, Japan; in 2015, she was selected "FEMtech Expert of the Month" by the Austrian Ministry for Transport, Innovation and Technology. She is currently working as tenure-track faculty at CISPA (Saarbrücken, Germany).

After completing her PhD "sub auspiciis praesidentis" (highest honor by the Austrian president), Johanna Ullrich became senior researcher at SBA Research and established a research group for network and CPS security. Since January 1st, 2020 she is a Key Researcher and SBA-K1 Area Manager (Area 1 Networked Systems Security) at SBA Research.

Recruiting

Efficiently staffing key positions in the field of IT research & consulting is one of the biggest challenges for research centers. The following measures have been implemented to interest people with the desired skill- and mindset:

- Stronger presence in relevant communities:
  - Events (sec4dev, ARES)
  - securepizza.club, CTFs
- Swift and transparent recruiting process
- Establishing and maintaining a database of suitable candidates
- Employer branding strategy
- Employee referral program
- Active sourcing of suitable candidates
- Improved overview of currently open positions on our website
- Enhanced social media presence
- Easy access to published papers and student activities

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt
wirtschaftsagentur wien
FWF Der Wissenschaftsfonds.
netidee
SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.

Controlling and Reporting @ SBA Research
Concept of an Efficient Multi-Stakeholder Controlling

Stefanie Schedlbauer

COMET Competence Centers for Excellent Technologies
www.ffg.at/comet
SBA Research

Structure of Financing and Costing in COMET SBA-K1

The overall structure of the financial report for COMET is shown in Figure 1. To achieve this, core principles which guarantee a sustainable business continuity must be implemented throughout the accounting system. Those principles are: transparency, credibility, consistency, and plausibility.
The main processes are included in every financial statement and presented on this poster. Based on this system, we are generating various recurring financial statements on different levels; furthermore it allows us to make fast ad-hoc statements according to specific requirements.

%	Financing from	annual
33,3	Federal funding	€ 1,378,667
16,6	Provincial funding	€ 689,333
5	Scientific InKind	€ 206,800
35	Company cash	€ 1,447,600
10	Company in-kind	€ 413,600
100	Sum	€ 4,136,000

%	Budget for	annual
43	Research projects directly cofunded by company partners	€ 1,780,000
24,6	SBA Research Initiatives (e.g. research, conferences, diaPRIMO)	€ 1,015,600
17,4	Overhead	€ 720,000
15	In-kind	€ 620,400
100	Sum	€ 4,136,000

Figure 1: Structure of financing and costing in COMET SBA-K1

Accounting: Data Processing

The first step in data processing is data acquisition. At SBA we have two main systems which share a large portion of the same data, i.e., personnel cost, material cost, and revenue.
Conventional accounting software rarely provides the flexible structure to ensure correct and fast financial overviews on a project level. Therefore, one of our systems has a special focus on the project level (ICM), and the other ensures that all legal requirements for the official annual financial statements are met (BMD).
Both systems are used for each report, at the very least to serve as first control instance for each other: if a report was created from one system, results from the other system must concur. The results in both systems have to be the same in relation to the shared database.
Data processing for every financial statement is shown in Figure 2. The individual instances and requirements are explained in Figure 3.

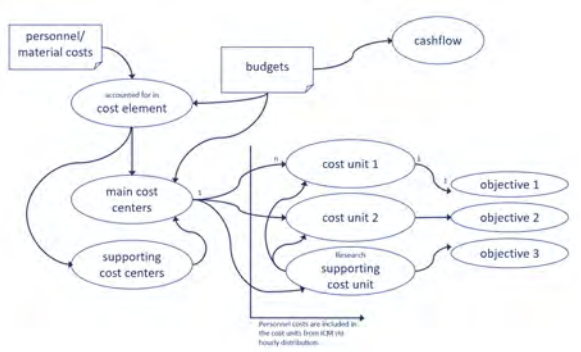


Figure 2: Financial statements and data process

Instance	Significance	Requirement
Cost element (account)	Nature of costs (e.g. personnel costs, travel expenses, etc.)	Definition of new accounts if needed
Main cost center (groups)	"Producing" areas/groups, meaning costs/revenue are directly attributable to project or product	Unique assignment of employees to areas/groups (and vice versa)
Supporting cost center (groups)	"Non-producing" areas/groups, meaning cost/revenue not directly attributable to project or product	Unique assignment of employees to areas/groups (and vice versa); distribution to all cost centers and units via secondary overhead cost calculation
Cost unit	Projects and products: material costs and income can be assigned to a project or product; for long-term and/or very large projects, they are shown in a project group with several cost units (e.g., COMET).	Personnel costs are included from ICM via hourly distribution. Overhead costs have to be distributed across all projects in a secondary calculation.
Supporting cost units	Special category for funded research projects, as some group costs shown in the main cost center are not in alignment with funding structures and cannot be assigned to a funded project.	Distribution via secondary overhead cost calculation
Strategic objectives	Assignment of cost units to objectives	As far as possible with n:1 relationship, exception: clear, temporally constant splitting
Budgets	Via combination of cost element/cost center	Inclusion in the planning control loop

Figure 3: Accounting structure of instances and requirements

Planning Control Loop

In order to connect all financial statements to target values and budgets we have established a planning control loop.

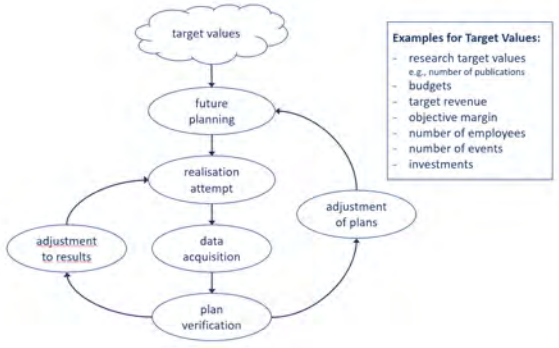


Figure 4: Planning control loop

Revision and Evaluation of Financial Statements

All our financial statements are regularly reviewed and evaluated. This includes:

i. Project statements, profitability analysis and cashflow are evaluated each quarter.
ii. COMET statements are submitted to the FFG twice per year and audited annually.
iii. Official annual financial statements, also audited externally.
iv. Governmental audits and project evaluations.

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.