Among Viruses, Trojans, and Backdoors

Fighting Malware in 2022

Marcus Botacin

[1]mfbotacin@inf.ufpr.br
mfbotacin@gmail.com
marcusbotacin.github.io

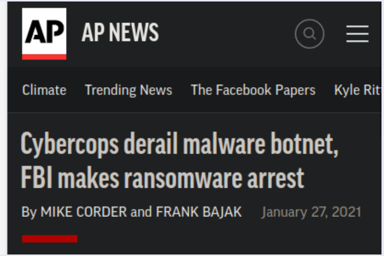
Who Am I?

- PhD. in Computer Science (2021) - Federal University of Paraná (UFPR), Brazil
  - Thesis: *“On the Malware Detection Problem: Challenges and new Approaches”*
- MSc. in Computer Science (2017) - University of Campinas (UNICAMP), Brazil
  - Dissertation: *“Hardware-Assisted Malware Analysis”*
- Computer Engineer (2015) - University of Campinas (UNICAMP), Brazil
  - Final Project: *“Malware detection via syscall patterns identification”*

Malware Detection
How have we been doing?

How have we been doing? (Malware Specifics)

The good side


AP NEWS

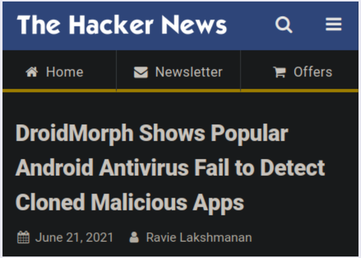
Climate · Trending News · The Facebook Papers · Kyle Rit

Cybercops derail malware botnet, FBI makes ransomware arrest

By MIKE CORDER and FRANK BAJAK January 27, 2021

Figure: **Source:** https://apnews.com/article/europe-malware-netherlands-coronavirus-pandemic-7de5f74120a968bd0a5bee3c57899fed

The bad side


The Hacker News

Home · Newsletter · Offers

DroidMorph Shows Popular Android Antivirus Fail to Detect Cloned Malicious Apps

June 21, 2021 · Ravie Lakshmanan

Figure: **Source:** https://thehackernews.com/2021/06/droidmorph-shows-popular-android.html

Malware Detection:
What have we been doing?

The State-of-the-art in Malware Detection & Prevention

Steps

1. Collection
2. Triage
3. Sandbox Analysis
4. Threat Intelligence
5. Endpoint Protection

Distributed Processing

- Collection

Cloud Processing

- Analysis and Intelligence steps

Limited Processing

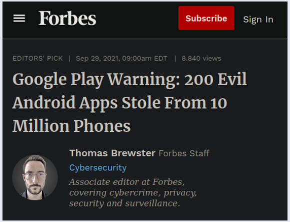
- Endpoint

Collection

How to find new malware samples?

- Searching “dark web” forums.
- Crawling software repositories.
- Leveraging honeypots.
- Checking spam traps.
- Downloading Malware repositories.
- Scrapping blocklists.

The result



≡ Forbes Subscribe Sign In

EDITORS' PICK | Sep 29, 2021, 09:00am EDT | 8,840 views

Google Play Warning: 200 Evil Android Apps Stole From 10 Million Phones

Thomas Brewster Forbes Staff
Cybersecurity
Associate editor at Forbes, covering cybercrime, privacy, security and surveillance.

Figure: **Source:** `https://www.forbes.com/sites/thomasbrewster/2021/09/29/google-play-warning-200-android-apps-stole-millions-from-10-million-phones/`

Triage

Why how many new malware samples?

- Variations from the same source code.

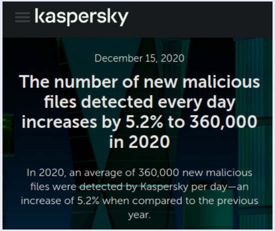
Implications

- Increase processing costs and response time.

How to solve this problem?

- Identify and cluster similar samples.

The Statistics



kaspersky

December 15, 2020

The number of new malicious files detected every day increases by 5.2% to 360,000 in 2020

In 2020, an average of 360,000 new malicious files were detected by Kaspersky per day—an increase of 5.2% when compared to the previous year.

Figure: **Source:** `https://www.kaspersky.com/about/press-releases/2020_the-number-of-new-malicious-files-detected-every-day-increases-by-52-to-360000-in-2020`

Sandbox Analysis

Goals

- Uncover hidden behaviors.

Method

- Trace sample execution.

Challenge

- Handle evasion attempts.

Solution 1

AUTOMATIC MALWARE ANALYSIS, DYNAMIC MALWARE ANALYSIS, MALWARE BEHAVIOUR, MULTISANDBOX, SANDBOX, VIRUSTOTAL SANDBOX, YOMI, YOROI

VirusTotal MultiSandbox += Yoroi: Yomi sandbox

Figure: `https://blog.virustotal.com/2019/05/virustotal-multisandbox-yoroi-yomi.html`

Solution 2

BEHAVIOR, MALWARE BEHAVIOR, MALWARE SANDBOX, MULTISANDBOX, SAND, SANDBOX, SNDBOX

VirusTotal MultiSandbox += SNDBOX

Figure: `https://blog.virustotal.com/2019/07/virustotal-multisandbox-sndbox.html`

Threat Intelligence

Goal

- Identify trends and predict attacks.

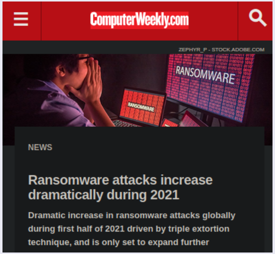
How?

- Data analytics over analyzed samples.

Challenges

- Look to a representative dataset.

We should look to:



ZEPHYR_P - STOCK.ADOBE.COM

NEWS

Ransomware attacks increase dramatically during 2021

Dramatic increase in ransomware attacks globally during first half of 2021 driven by triple extortion technique, and is only set to expand further

Figure: **Source:** `https://www.computerweekly.com/news/252504676/Ransomware-attacks-increase-dramatically-during-2021`

Endpoint Protection

Goal

- Protect customers in their machines.

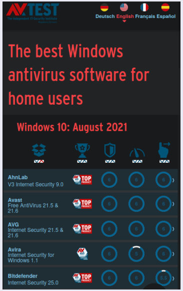
How?

- Moving the **viable** analyses to the endpoint.

Challenges

- Performance and usability constraints.

Is there a "best"?



The best Windows antivirus software for home users

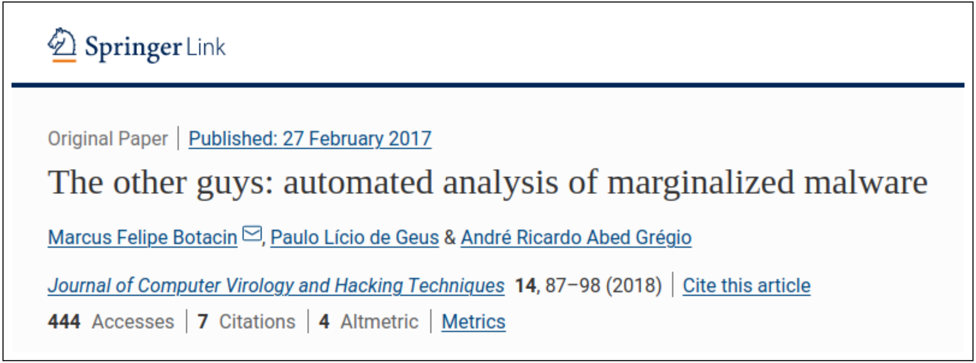
Windows 10: August 2021

AhnLab V3 Internet Security 9.0	TOP PRODUCT	6	6	6
Avast Free AntiVirus 21.5 & 21.6	TOP PRODUCT	6	6	6
AVG Internet Security 21.5 & 21.6	TOP PRODUCT	6	6	6
Avira Internet Security for Windows 1.1		6	5	6
Bitdefender Internet Security 25.0	TOP PRODUCT	6	6	5.5

Figure: **Source:** `https://www.av-test.org/en/antivirus/home-windows/`

Enhancing Malware Tracing

Publication



SpringerLink

Original Paper | Published: 27 February 2017

The other guys: automated analysis of marginalized malware

Marcus Felipe Botacin, Paulo Lício de Geus & André Ricardo Abed Grégio

Journal of Computer Virology and Hacking Techniques **14**, 87–98 (2018) | Cite this article

444 Accesses | **7** Citations | **4** Altmetric | Metrics

Figure: **Link:** `https://link.springer.com/article/10.1007/s11416-017-0292-8`

Software-based Sandbox

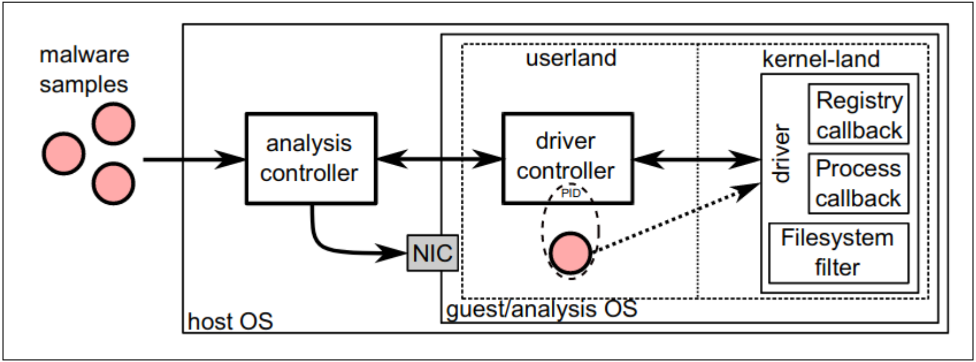
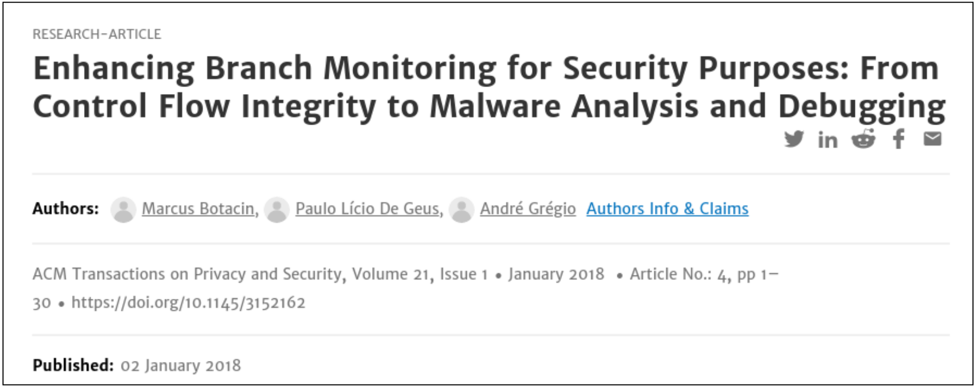


Figure: **System Architecture.** Analysis VMs.

Publication



RESEARCH-ARTICLE

Enhancing Branch Monitoring for Security Purposes: From Control Flow Integrity to Malware Analysis and Debugging

Authors: Marcus Botacin, Paulo Lício De Geus, André Grégio Authors Info & Claims

ACM Transactions on Privacy and Security, Volume 21, Issue 1 • January 2018 • Article No.: 4, pp 1–30 • https://doi.org/10.1145/3152162

Published: 02 January 2018

Figure: **Link:** `https://dl.acm.org/doi/10.1145/3152162`

Hardware-based Sandbox

Monitoring Steps

1. Software executes a branch.
2. Processor stores branch address in memory page.
3. Processor raises an interrupt.
4. Kernel handles interrupt.
5. Kernel sends data to userland.
6. Userland introspects into this data.

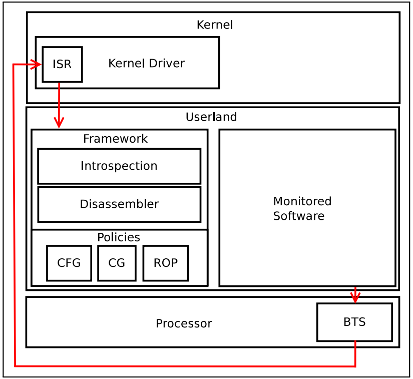


Figure: **System Architecture.**

Key Insight: Branches define basic blocks

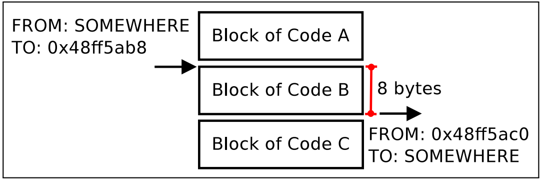
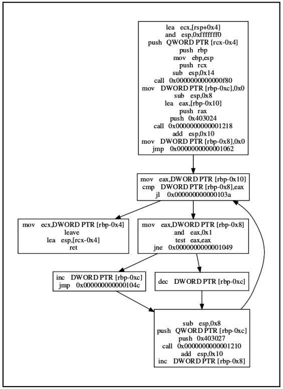


Figure: **Identified branches and basic blocks.**



```
lea   ecx,[rsp+0x4]
and   esp,0xfffffff0
push  QWORD PTR [rcx-0x4]
push  rbp
mov   ebp,esp
push  rcx
sub   esp,0x14
call  0x0000000000000f80
mov   DWORD PTR [rbp-0xc],0x0
sub   esp,0x8
lea   eax,[rbp-0x10]
push  rax
push  0x403024
call  0x0000000000001218
add   esp,0x10
mov   DWORD PTR [rbp-0x8],0x0
jmp   0x0000000000001062
```

```
mov   eax,DWORD PTR [rbp-0x10]
cmp   DWORD PTR [rbp-0x8],eax
jl    0x000000000000103a
```

```
mov   ecx,DWORD PTR [rbp-0x4]
leave
lea   esp,[rcx-0x4]
ret
```

```
mov   eax,DWORD PTR [rbp-0x8]
and   eax,0x1
test  eax,eax
jne   0x0000000000001049
```

```
inc   DWORD PTR [rbp-0xc]
jmp   0x000000000000104c
```

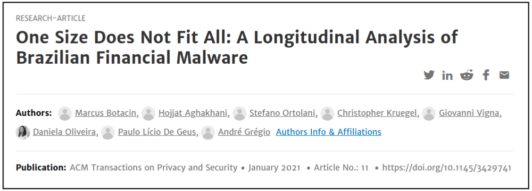
```
dec   DWORD PTR [rbp-0xc]
```

```
sub   esp,0x8
push  QWORD PTR [rbp-0xc]
push  0x403027
call  0x0000000000001210
add   esp,0x10
inc   DWORD PTR [rbp-0x8]
```

Figure: **CFG Reconstruction.**

From Tracing to Threat Intelligence

Publications



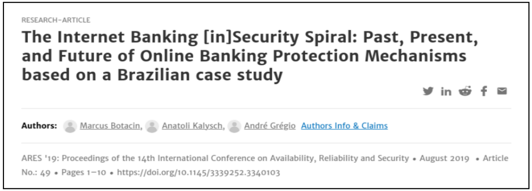
RESEARCH-ARTICLE

One Size Does Not Fit All: A Longitudinal Analysis of Brazilian Financial Malware

Authors: Marcus Botacin, Hojjat Aghakhani, Stefano Ortolani, Christopher Kruegel, Giovanni Vigna, Daniela Oliveira, Paulo Lício De Geus, André Grégio Authors Info & Affiliations

Publication: ACM Transactions on Privacy and Security • January 2021 • Article No.: 11 • https://doi.org/10.1145/3429741

Figure: **Link:** `https://dl.acm.org/doi/10.1145/3429741`



RESEARCH-ARTICLE

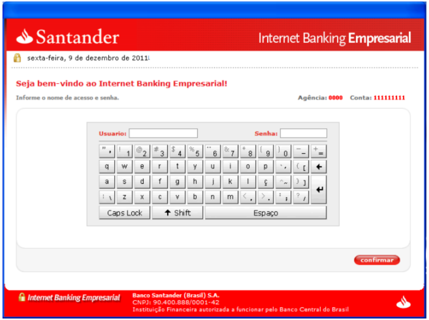
The Internet Banking [in]Security Spiral: Past, Present, and Future of Online Banking Protection Mechanisms based on a Brazilian case study

Authors: Marcus Botacin, Anatoli Kalysch, André Grégio Authors Info & Claims

ARES '19: Proceedings of the 14th International Conference on Availability, Reliability and Security • August 2019 • Article No.: 49 • Pages 1–10 • https://doi.org/10.1145/3339252.3340103

Figure: **Link:** `https://dl.acm.org/doi/10.1145/3339252.3340103`

Brazilian Financial Malware on Desktop



Santander Internet Banking Empresarial

sexta-feira, 9 de dezembro de 2011

Seja bem-vindo ao Internet Banking Empresarial!

Informe o nome de acesso e senha.

Agência: 0000 Conta: 111111111

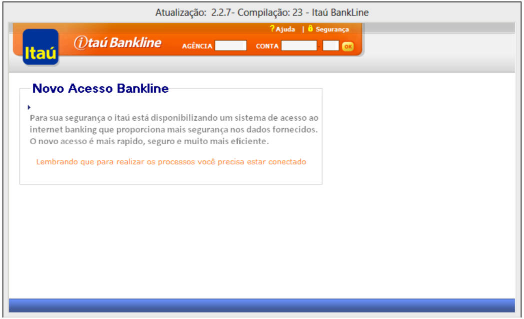
Usuario:

Senha:

Internet Banking Empresarial

Banco Santander (Brasil) S.A.
CNPJ: 90.400.888/0001-42
Instituição Financeira autorizada a funcionar pelo Banco Central do Brasil

Figure: **Passive Banker Malware for Santander bank** waiting for user's credential input.



Atualização: 2.2.7- Compilação: 23 - Itaú BankLine

AGÊNCIA

CONTA

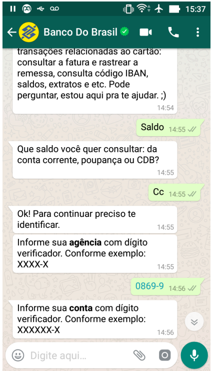
Novo Acesso Bankline

Para sua segurança o itaú está disponibilizando um sistema de acesso ao internet banking que proporciona mais segurança nos dados fornecidos. O novo acesso é mais rapido, seguro e muito mais eficiente.

Lembrando que para realizar os processos você precisa estar conectado

Figure: **Passive Banker Malware for Itaú bank** waiting for user's credential input.

Brazilian Financial Malware on Mobile



transações relacionadas ao cartão: consultar a fatura e rastrear a remessa, consulta código IBAN, saldos, extratos e etc. Pode perguntar, estou aqui pra te ajudar. ;) 14:54

Saldo 14:55

Que saldo você quer consultar: da conta corrente, poupança ou CDB? 14:55

Cc 14:55

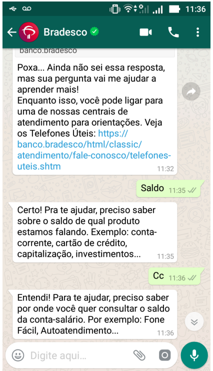
Ok! Para continuar preciso te identificar. 14:55

Informe sua **agência** com dígito verificador. Conforme exemplo: XXXX-X 14:55

0869-9 14:56

Informe sua **conta** com dígito verificador. Conforme exemplo: XXXXXX-X 14:56

Figure: **BB's Whatsapp chatbot.**



Poxa... Ainda não sei essa resposta, mas sua pergunta vai me ajudar a aprender mais!
Enquanto isso, você pode ligar para uma de nossas centrais de atendimento para orientações. Veja os Telefones Úteis: https://banco.bradesco/html/classic/atendimento/fale-conosco/telefones-uteis.shtm 11:32

Saldo 11:35

Certo! Pra te ajudar, preciso saber sobre o saldo de qual produto estamos falando. Exemplo: conta-corrente, cartão de crédito, capitalização, investimentos... 11:35

Cc 11:36

Entendi! Para te ajudar, preciso saber por onde você quer consultar o saldo da conta-salário. Por exemplo: Fone Fácil, Autoatendimento... 11:36

Figure: **Bradesco's Whatsapp chatbot.**

More about Brazilian Malware

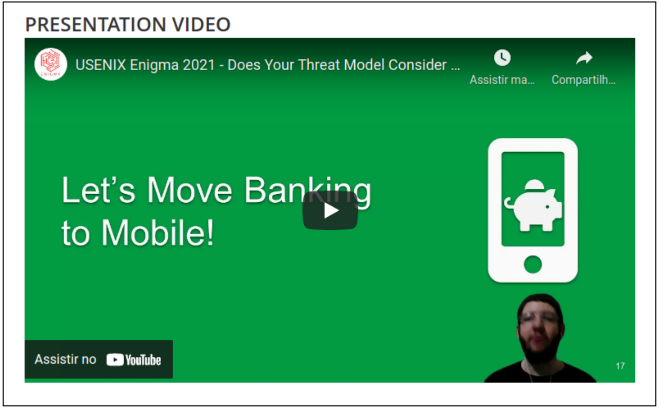
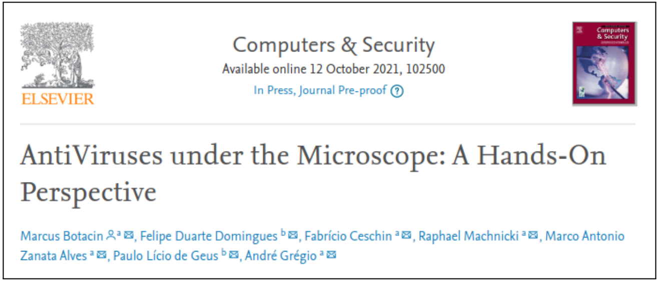

PRESENTATION VIDEO
USENIX Enigma 2021 - Does Your Threat Model Consider ...
Assistir ma...
Compartilh...
Let's Move Banking to Mobile!
Assistir no YouTube
17

Figure: **Link:** `https://www.usenix.org/conference/enigma2021/presentation/botacin`

From Threat Intelligence to Endpoint Protection

Publication



ELSEVIER

Computers & Security
Available online 12 October 2021, 102500
In Press, Journal Pre-proof

AntiViruses under the Microscope: A Hands-On Perspective

Marcus Botacin [a], Felipe Duarte Domingues [b], Fabrício Ceschin [a], Raphael Machnicki [a], Marco Antonio Zanata Alves [a], Paulo Lício de Geus [b], André Grégio [a]

Figure: **Link:** `https://www.sciencedirect.com/science/article/pii/S0167404821003242`

Drawback: Real-time monitoring performance penalty

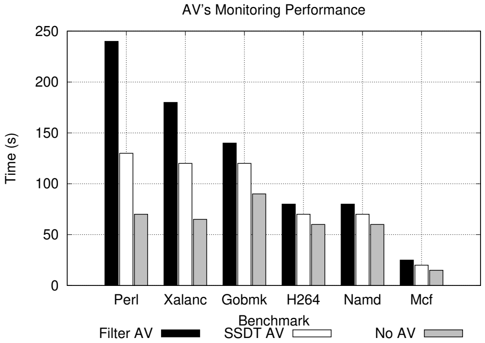


Figure: **AV Monitoring Performance.**

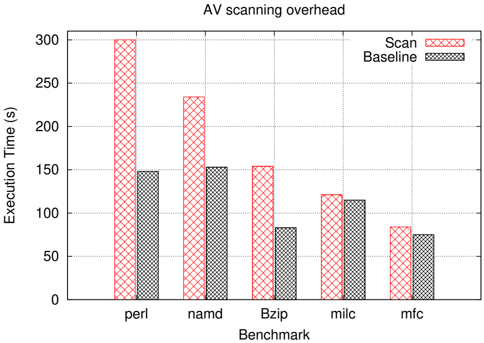


Figure: **In-memory AV scans worst-case and best-case performance penalties.**

Publication

The AV says: Your Hardware Definitions Were Updated!

Publisher: IEEE

Marcus Botacin ; Lucas Galante ; Fabricio Ceschin ; Paulo C. Santos ; Luigi Carro ; Paulo de Geus ; André Grégio ; Marco A. Z. ... **All Authors**

Figure: **Link:** `https://ieeexplore.ieee.org/document/9034972`

SMC-Aware Processor

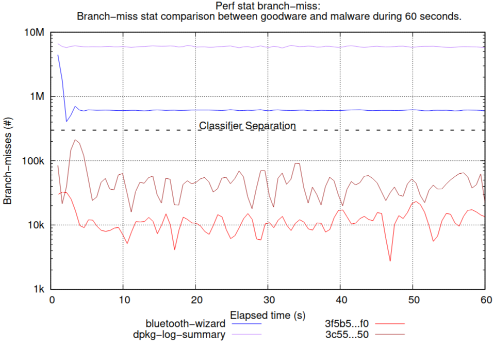


Figure: **Sample Profiling.**

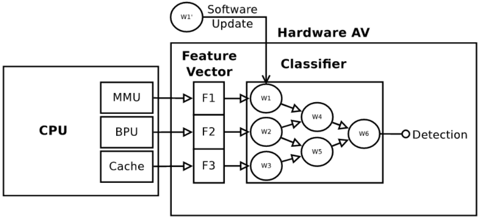
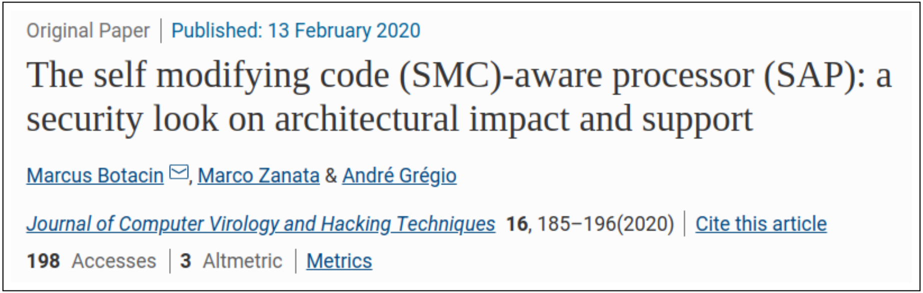


Figure: **System Overview.**

Publication



Original Paper | Published: 13 February 2020

The self modifying code (SMC)-aware processor (SAP): a security look on architectural impact and support

Marcus Botacin, Marco Zanata & André Grégio

Journal of Computer Virology and Hacking Techniques **16**, 185–196(2020) | Cite this article

198 Accesses | **3** Altmetric | Metrics

Figure: **Link:** `https://link.springer.com/article/10.1007/s11416-020-00348-w`

SMC-Aware Processor

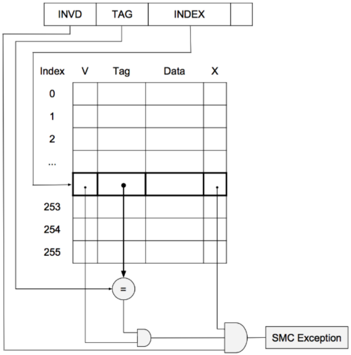
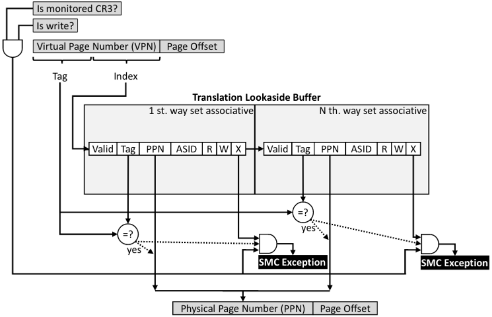


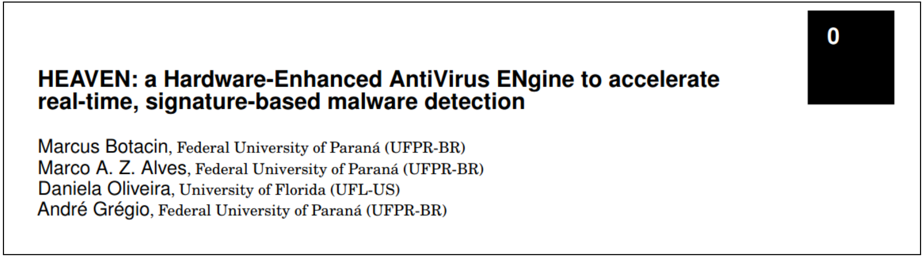
Figure: **Modified Cache.**


Is monitored CR3?
Is write?
Virtual Page Number (VPN)
Page Offset
Tag
Index
Translation Lookaside Buffer
1 st. way set associative
N th. way set associative
Valid
Tag
PPN
ASID
R
W
X
Valid
Tag
PPN
ASID
R
W
X
=?
yes
=?
yes
SMC Exception
SMC Exception
Physical Page Number (PPN)
Page Offset

MMU-based SMC detection mechanism.

Figure: **Modified MMU.**

Publication



HEAVEN: a Hardware-Enhanced AntiVirus ENgine to accelerate real-time, signature-based malware detection

Marcus Botacin, Federal University of Paraná (UFPR-BR)
Marco A. Z. Alves, Federal University of Paraná (UFPR-BR)
Daniela Oliveira, University of Florida (UFL-US)
André Grégio, Federal University of Paraná (UFPR-BR)

Figure: **Link:** `https://www.sciencedirect.com/science/article/abs/pii/S0957417422004882`

A first idea: Hardware features as signatures

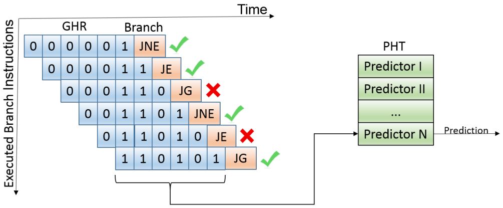


Figure: **Two-level branch predictor.** A sequence window of taken (1) and not-taken (0) branches is stored in the Global History Register (GHR).

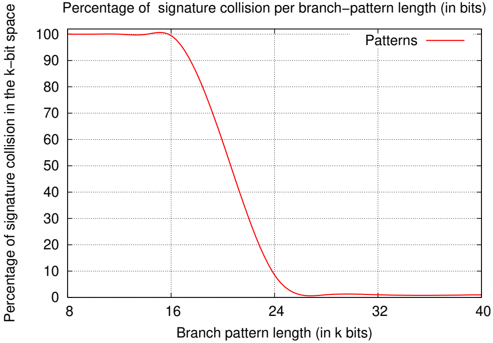
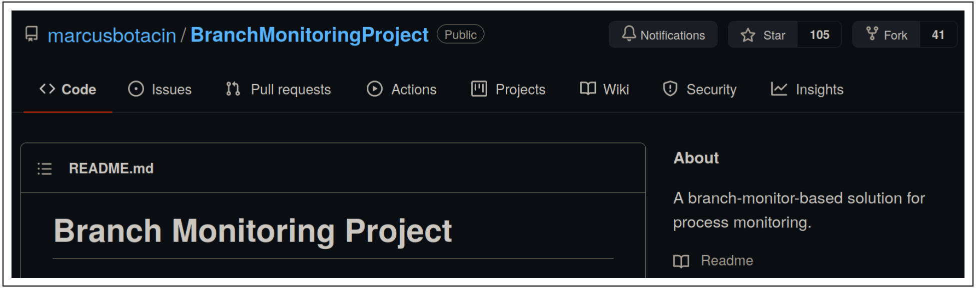


Figure: **Branch patterns coverage.**

Solutions Availability

Code: The BranchMonitoring Project



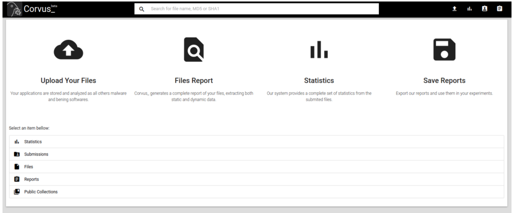
Branch Monitoring Project

About

A branch-monitor-based solution for process monitoring.

Figure: **Link:** `https://github.com/marcusbotacin/BranchMonitoringProject`

Service: Corvus Platform



Upload Your Files

Your applications are stored and analyzed as all others malware and bening softwares.

Files Report

Corvus_ generates a complete report of your files, extracting both static and dynamic data.

Statistics

Our system provides a complete set of statistics from the submited files.

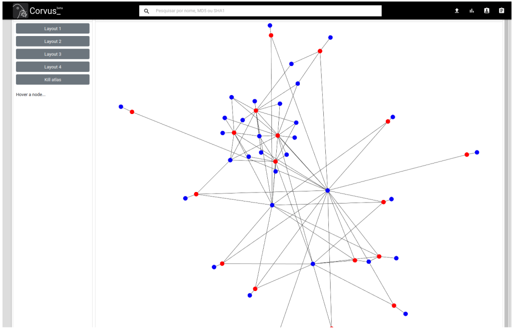
Save Reports

Export our reports and use them in your experiments.

Select an item bellow:

- Statistics
- Submissions
- Files
- Reports
- Public Collections

Figure: **Link:** `corvus.inf.ufpr.br`

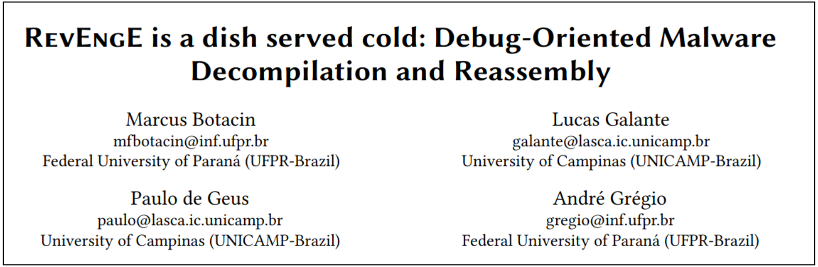


Hover a node...

Figure: **Corvus'** Threat Intelligence.

Current Research:
Malware Decompilation

Publication



REVENGE is a dish served cold: Debug-Oriented Malware Decompilation and Reassembly

Marcus Botacin
mfbotacin@inf.ufpr.br
Federal University of Paraná (UFPR-Brazil)

Lucas Galante
galante@lasca.ic.unicamp.br
University of Campinas (UNICAMP-Brazil)

Paulo de Geus
paulo@lasca.ic.unicamp.br
University of Campinas (UNICAMP-Brazil)

André Grégio
gregio@inf.ufpr.br
Federal University of Paraná (UFPR-Brazil)

Figure: **Link:** `https://dl.acm.org/doi/10.1145/3375894.3375895`

Decompilation Execution Example 1

Data Extraction

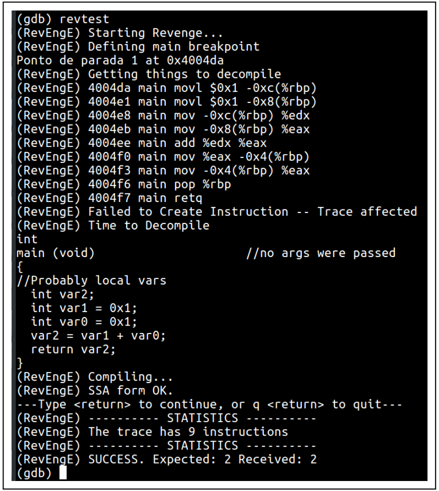
- Debugging with GDB.

Decompilation

- Lifting with Python.

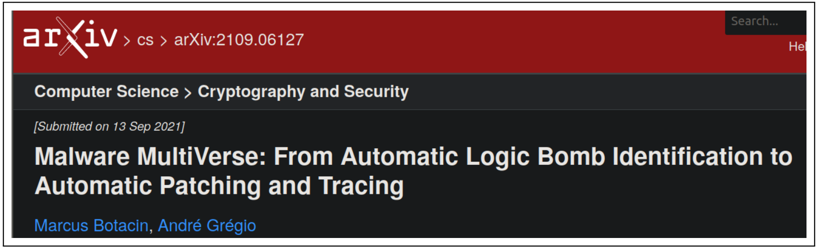
Recompilation

- Using GCC.



```
(gdb) revtest
(RevEngE) Starting Revenge...
(RevEngE) Defining main breakpoint
Ponto de parada 1 at 0x4004da
(RevEngE) Getting things to decompile
(RevEngE) 4004da main movl $0x1 -0xc(%rbp)
(RevEngE) 4004e1 main movl $0x1 -0x8(%rbp)
(RevEngE) 4004e8 main mov -0xc(%rbp) %edx
(RevEngE) 4004eb main mov -0x8(%rbp) %eax
(RevEngE) 4004ee main add %edx %eax
(RevEngE) 4004f0 main mov %eax -0x4(%rbp)
(RevEngE) 4004f3 main mov -0x4(%rbp) %eax
(RevEngE) 4004f6 main pop %rbp
(RevEngE) 4004f7 main retq
(RevEngE) Failed to Create Instruction -- Trace affected
(RevEngE) Time to Decompile
int
main (void)                     //no args were passed
{
//Probably local vars
  int var2;
  int var1 = 0x1;
  int var0 = 0x1;
  var2 = var1 + var0;
  return var2;
}
(RevEngE) Compiling...
(RevEngE) SSA form OK.
---Type <return> to continue, or q <return> to quit---
(RevEngE) ---------- STATISTICS ----------
(RevEngE) The trace has 9 instructions
(RevEngE) ---------- STATISTICS ----------
(RevEngE) SUCCESS. Expected: 2 Received: 2
(gdb)
```

Publication



Computer Science > Cryptography and Security

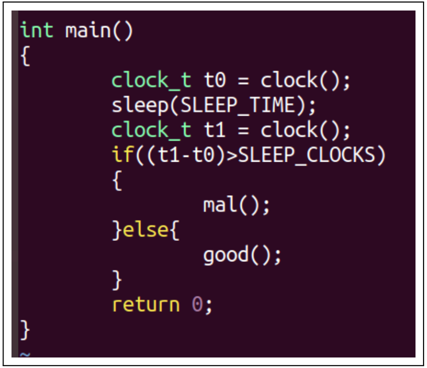
[Submitted on 13 Sep 2021]

Malware MultiVerse: From Automatic Logic Bomb Identification to Automatic Patching and Tracing

Marcus Botacin, André Grégio

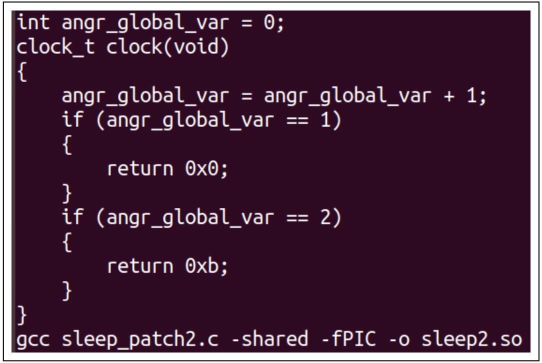
Figure: **Link:** `https://arxiv.org/abs/2109.06127`

Decompilation Execution Example 1



```
int main()
{
        clock_t t0 = clock();
        sleep(SLEEP_TIME);
        clock_t t1 = clock();
        if((t1-t0)>SLEEP_CLOCKS)
        {
                mal();
        }else{
                good();
        }
        return 0;
}
~
```

Figure: **Malware Source-Code.**

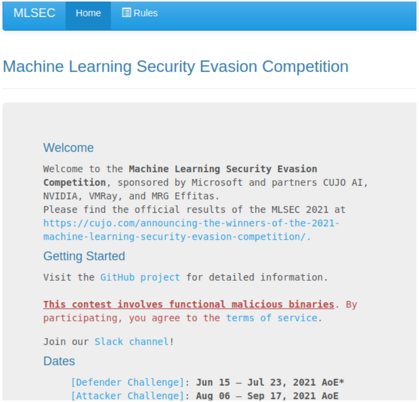


```
int angr_global_var = 0;
clock_t clock(void)
{
    angr_global_var = angr_global_var + 1;
    if (angr_global_var == 1)
    {
        return 0x0;
    }
    if (angr_global_var == 2)
    {
        return 0xb;
    }
}
gcc sleep_patch2.c -shared -fPIC -o sleep2.so
```

Figure: **Generated Patch.**

Machine Learning:
The Latest Trend

Malware Evasion Competition



Machine Learning Security Evasion Competition

Welcome

Welcome to the **Machine Learning Security Evasion Competition**, sponsored by Microsoft and partners CUJO AI, NVIDIA, VMRay, and MRG Effitas.
Please find the official results of the MLSEC 2021 at https://cujo.com/announcing-the-winners-of-the-2021-machine-learning-security-evasion-competition/.

Getting Started

Visit the GitHub project for detailed information.

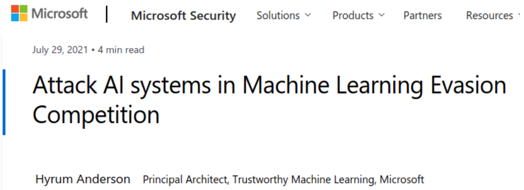
This contest involves functional malicious binaries. By participating, you agree to the terms of service.

Join our Slack channel!

Dates

[Defender Challenge]: **Jun 15 – Jul 23, 2021 AoE***
[Attacker Challenge]: **Aug 06 – Sep 17, 2021 AoE**

Figure: **Source:** `mlsec.io`



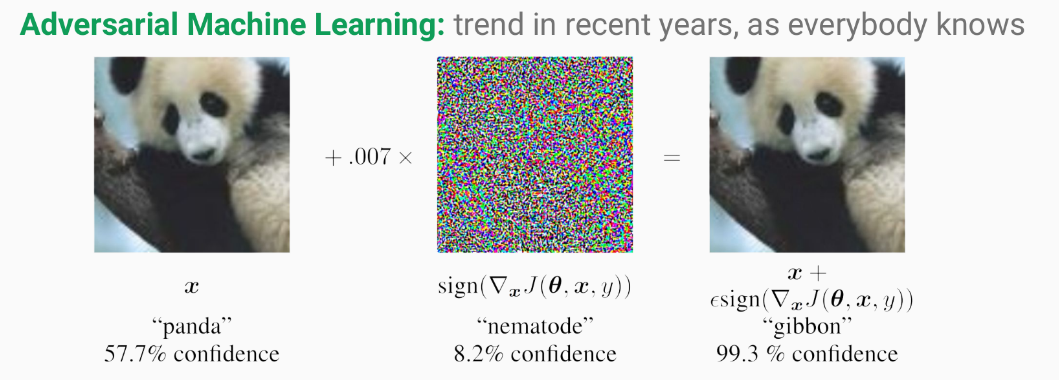
July 29, 2021 • 4 min read

Attack AI systems in Machine Learning Evasion Competition

Hyrum Anderson Principal Architect, Trustworthy Machine Learning, Microsoft

Figure: **Source:** `https://www.microsoft.com/security/blog/2021/07/29/attack-ai-systems-in-machine-learning-evasion-competition/`

Adversarial Machine Learning



Adversarial Machine Learning: trend in recent years, as everybody knows

$\boldsymbol{x}$
"panda"
57.7% confidence
$+ .007 \times$
$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$
"nematode"
8.2% confidence
$=$
$\boldsymbol{x} + \epsilon\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$
"gibbon"
99.3 % confidence

Figure: **Source:** `https://github.com/marcusbotacin/Talks/tree/master/Waikato`

Adversarial Malware

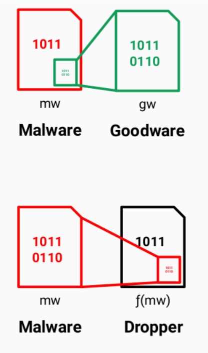


Figure: **Dropper Strategy.**

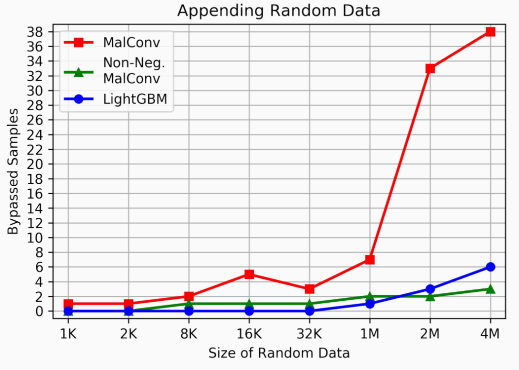
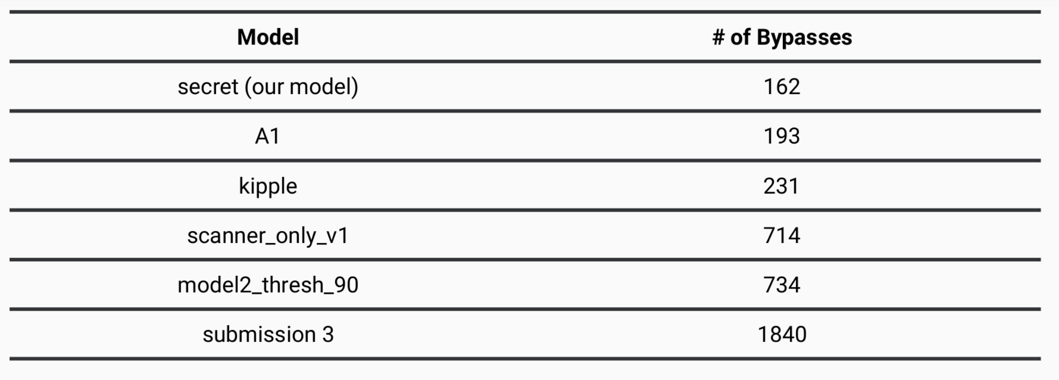


Figure: **Data Appendix Result.**

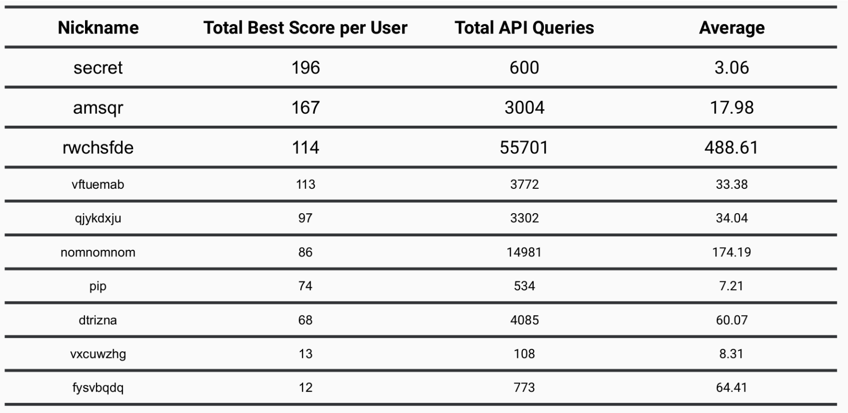
Challenge Results



Model	# of Bypasses
secret (our model)	162
A1	193
kipple	231
scanner_only_v1	714
model2_thresh_90	734
submission 3	1840

Figure: **Defenders Challenge.**

Challenge Results



Nickname	Total Best Score per User	Total API Queries	Average
secret	196	600	3.06
amsqr	167	3004	17.98
rwchsfde	114	55701	488.61
vftuemab	113	3772	33.38
qjykdxju	97	3302	34.04
nomnomnom	86	14981	174.19
pip	74	534	7.21
dtrizna	68	4085	60.07
vxcuwzhg	13	108	8.31
fysvbqdq	12	773	64.41

Figure: **Attackers Challenge.**

What's Next?

Thanks!

Questions? Comments?

@MarcusBotacin
mfbotacin@inf.ufpr.br
mfbotacin@gmail.com
marcusbotacin.github.io
corvus.inf.ufpr.br