Matrioska: A User-Centric Defense Against Virtualization-Based Repackaging Malware on Android

Samuele Doria samueledoria@gmail.com



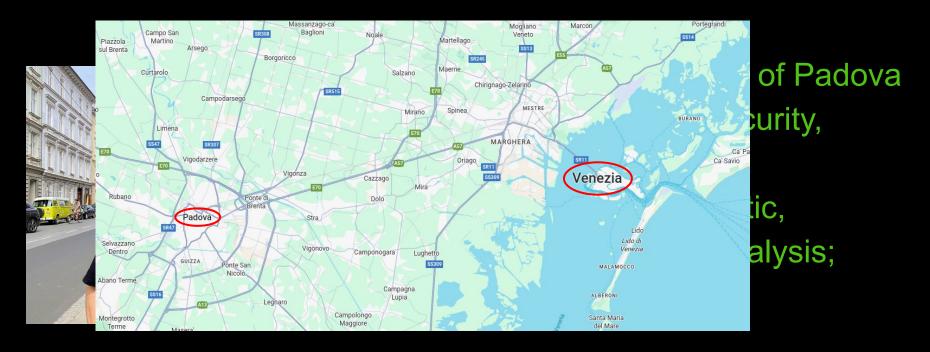


\$ whoami



- PhD Student @ University of Padova
- Researching Software Security, focusing on Android
 - Malware detection; static,
 dynamic and hybrid analysis;
 etc...

\$ whoami



\$ cat talk.txt

R+R: Matrioska: A User-Centric Defense Against Virtualization-Based Repackaging Malware on Android

Simone Zerbini
Department of Mathematics
University of Padua
Padua, Italy
simone.zerbini@unipd.it

Samuele Doria
Department of Mathematics
University of Padua
Padua, Italy
sdoria@math.unipd.it

Primal Wijesekera International Computer Science Institute Berkeley, USA primalw@icsi.berkeley.edu

Serge Egelman
International Computer Science Institute /
University of California Berkeley
Berkeley, USA
egelman@cs.berkeley.edu

Eleonora Losiouk
Department of Mathematics
University of Padua
Padua, Italy
eleonora.losiouk@unipd.it





Why care about this talk

- Listen to some cool academic research (unbiased opinion :))
- You like Android, malware, and dynamic analysis
- You want to know what virtualization on Android is
 - spoiler: it's different from traditional virtualization

Why it's important

- Recently, new malware samples using virtualization have been found [1]
- Zimperium discovered 29 new malicious apps using virtualization containers
 - Hundreds of thousands of installs across third-party markets

Why it's important

- The malware includes:
 - A user-facing app (e.g., fake banking/utility app)
 - An embedded plugin APK that's silently loaded
- Malware behavior happens inside the plugin, out of reach of most defenses

Why it's important

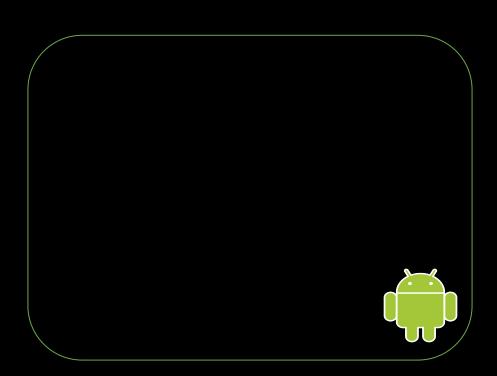
- Plugin apps downloaded at runtime from C2 servers
- Threats: Credential theft, overlay attacks, ad fraud, spyware, and more
- Detection Challenges: No APK installation, hides inside a Virtual Environment

But, what is a Virtualization-Based Repackaging malware?

 Repackaging malware is a modified version of a legitimate app where malicious code is injected.

 Virtualization-Based Repackaging (VBR) malware takes advantage of virtualization to deploy the attacks.

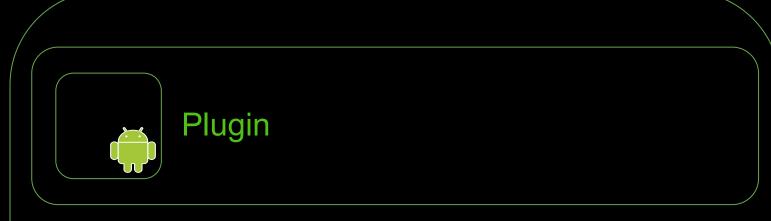
But, what is Android's Virtualization?



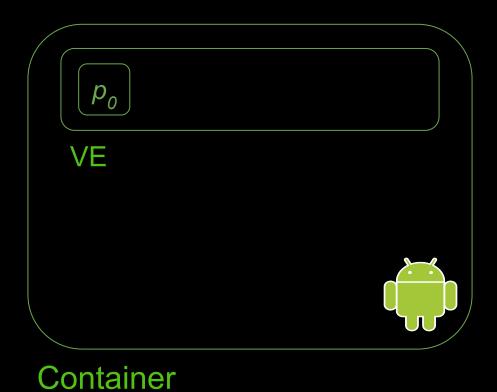


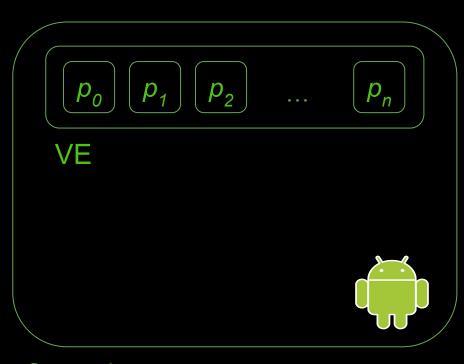


Container



Virtual Environment (VE)





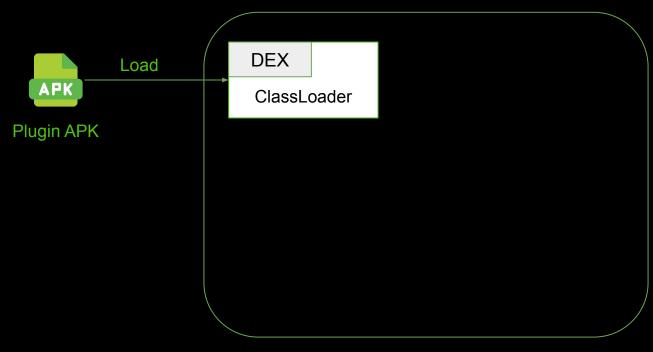
Characteristics:

- sharing the UID and the permissions
- breaking app isolation
- p_i = plugin(s)

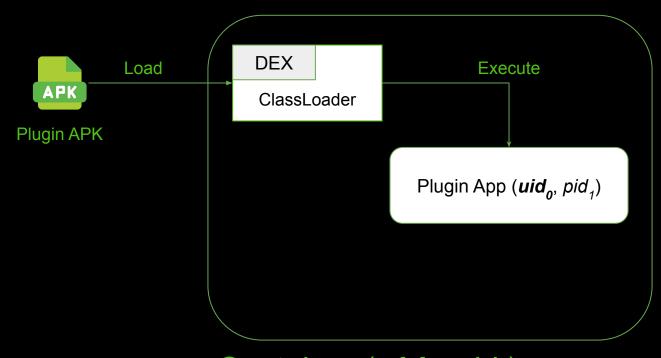
Container



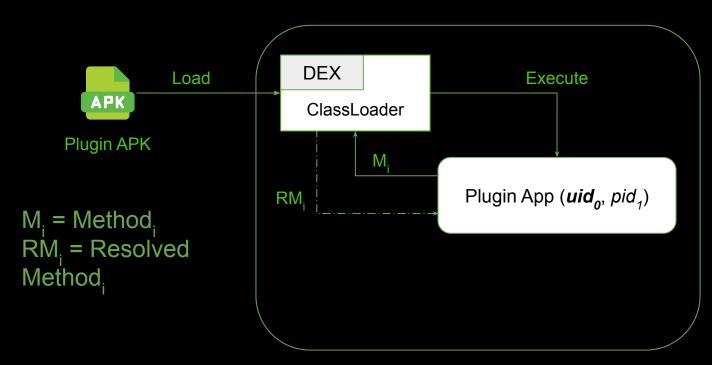




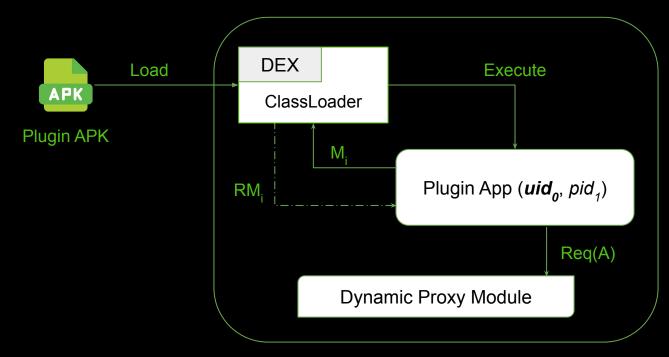




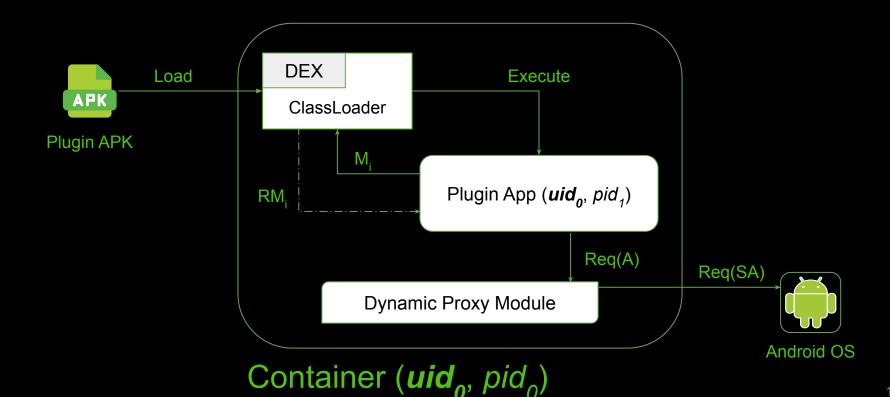




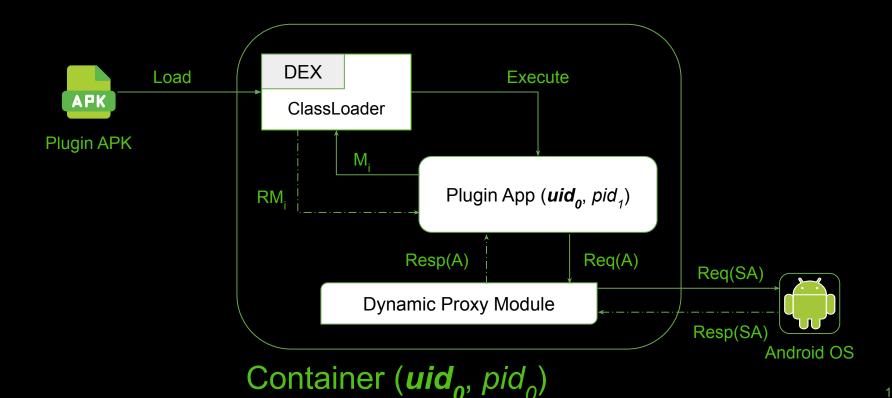


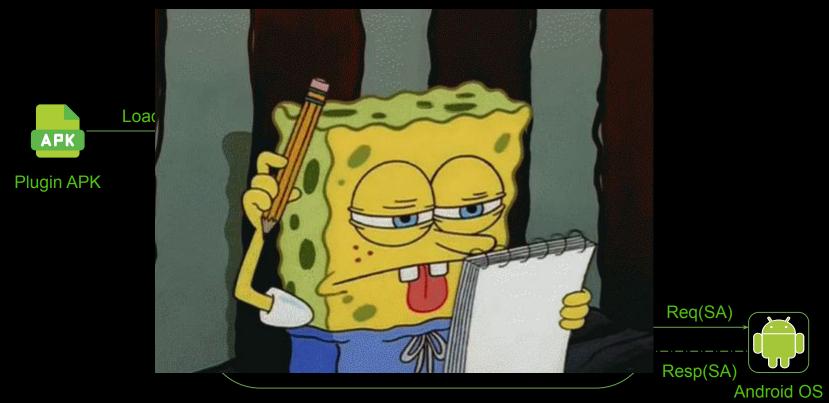


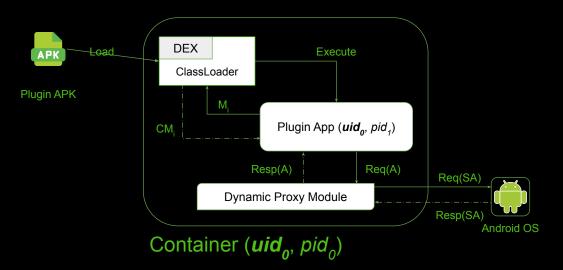




11/58







- Container and plugin have the same **UID**
- The container has to declare all permissions
- The container has to manage the life cycle of all the plugin's components

Virtualization Frameworks

- Many virtualization frameworks are open-source (e.g., VirtualApp [2], DroidPlugin [3])
- Developers quickly found out they could also add hooking capabilities without needing root privileges
 - YAHFA [4]

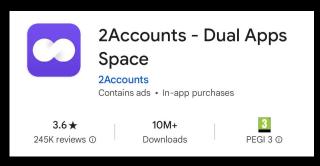
- [2] https://github.com/asLody/VirtualApp
- [3] https://github.com/DroidPluginTeam/DroidPlugin
- [4] https://github.com/PAGalaxyLab/YAHFA

Multiple purposes:

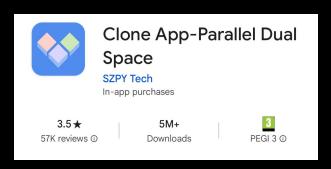
- multiple instances of the same app
- sandboxing malware

but also...

deploying malware

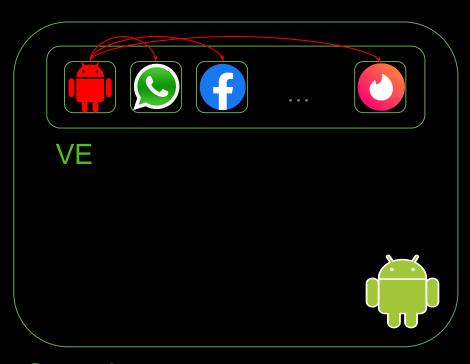


Source: Google Play Store



Source: Google Play Store

Malicious plugin



Malicious plugin:

- abuses the inherited permissions from the container
- steals data
- injects code

Container

Virtualization-Based Repackaging (VBR) malware



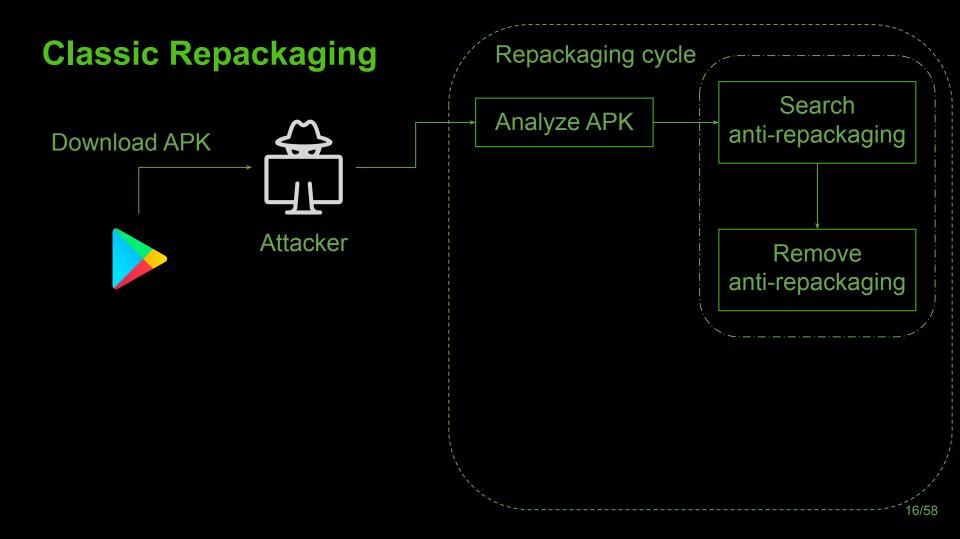
Malicious Container:

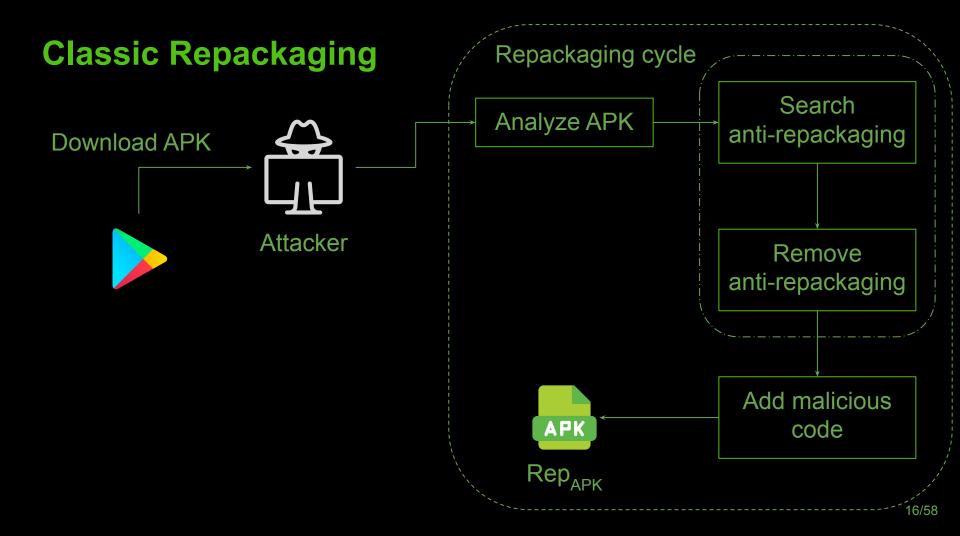
- controls the plugins' execution
- steals credentials and private information
- injects advertisement

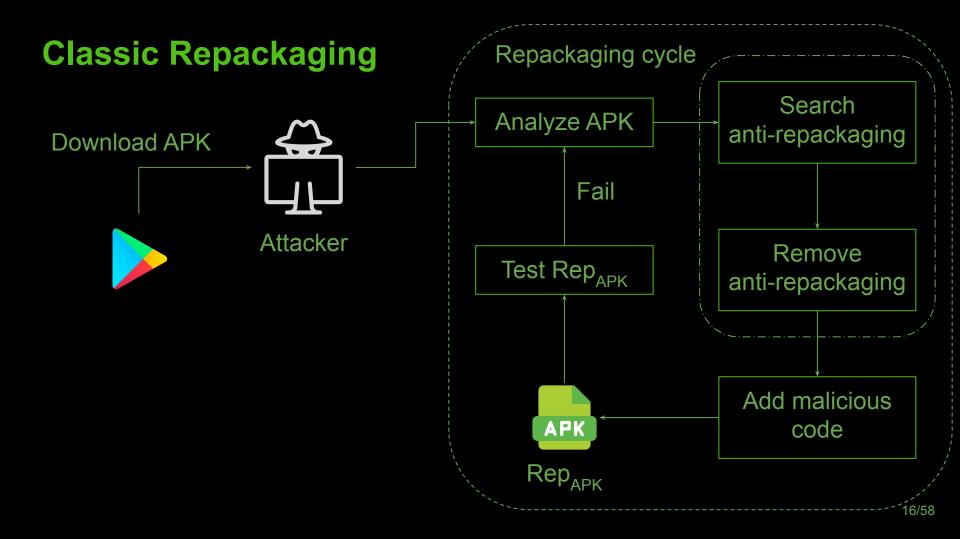
Container

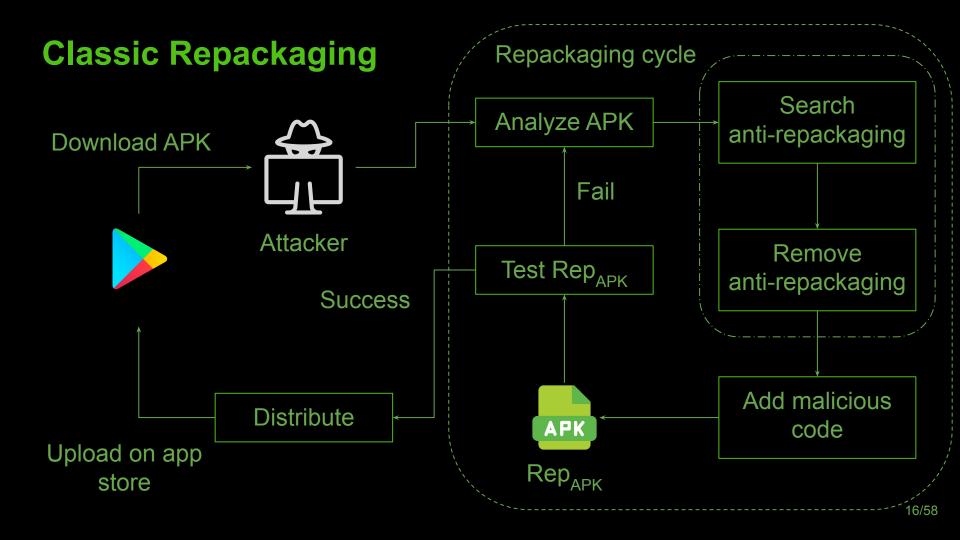
Classic Repackaging



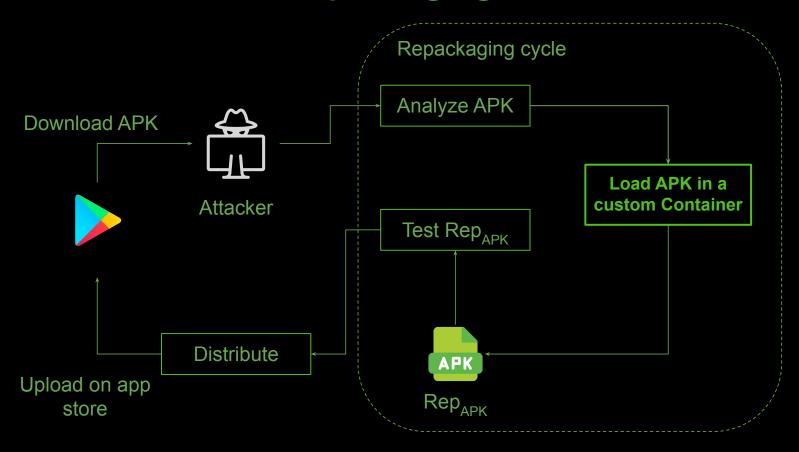








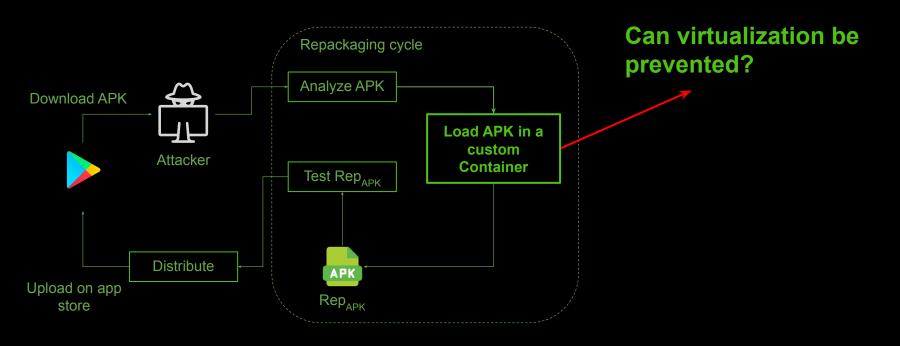
Virtualization-Based Repackaging



Repackaging Compared

	Classic Repackaging	Virtualization-Based Repackaging
Malicious Payload Location	Injected directly into the APK	Part of the custom Container
Signature Checks	App's signature is modified	Plugin remain unmodified
Visibility	Visible as an installed app	Plugin can remain hidden
Legitimate use	Rare	Common in dual-app tools

Virtualization-Based Repackaging



Can Virtualization-Based Repackaging be detected?

State-of-the-art

- VAHunt [5] is the state-of-the-art static analysis tool for VBR malware
- Authors reported to have found 71,303 VBR malware in a dataset of 139K virtualization-based apps

State-of-the-art

- Four anti-virtualization libraries have been previously proposed
- They can be embedded in the apps and recognize the presence of virtualization

Are they effective and adopted among the most downloaded apps?

State-of-the-art

- Marvel [6] prevents repackaging attacks by forcing the protected apps to run only in a trusted container
- The apps are protected by removing portions of code, that are re-introduced by the trusted container

Is Marvel enough to protect against VBR malware?

Motivation

Do we have to be worried about VBR malware?

 VAHunt found <u>71,303 VBR malware</u> over a dataset of 139K virtualization malware

Driving Research Questions

RQ1: How effective are existing defence mechanisms against VBR malware?

RQ2: How prevalent are the current detection and defence mechanisms against VBR malwares?

- We tested the four anti-virtualization libraries, comprising a total of 19 checks
- These are libraries embedded in apps to detect if they are running inside a Virtual Environment (VE).

 We crafted a container and used YAHFA to hook the checks and alter their results

Our crafted container is able to bypass all of them

- Being able to access more permissions than the ones declared
 - We customized the container to conservatively request the plugin's permissions

- Querying package name and components through the PackageManager
 - Hook the APIs and modify the response

- Checking the UID and other running processes with the same one
 - Hook the getRunningServices API and remove the container listing

- Throw an exception and analyzing the stack trace
 - Hook the getStackTrace API and remove references to the virtualization framework

Video Demo

```
IN THE STATE AND THE TOTAL PRODUCTIONS OF THE STATE AND ADDRESS OF THE STATE ADDRESS OF THE STATE ADDRESS OF THE STATE AND ADDRESS OF THE STATE ADDRESS OF THE STATE ADDRESS OF THE STATE ADDRESS OF THE STATE ADDRESS OF THE

    contents office Epidemicrosis use our proper printed line nonclinoshed. Annalsomorphism is view of

  on to Paper with 1988
the per parties with a few times a second business insured to continuous come indicately a second to
  IK IN EVEL 2001 VIVI VIVI V. Ovjewani seksembandahan sakuntan Kurinta meluka dana seksa
                                                                                                                                                                                                                                                                                                                                                                                     EXHIBITAL perfective/object in Fider
  IN SECURITY THE COST CALL THE SECURITY VICTORIAN STREET, THESE WHEN AND THE
  on the second state of the second of the
  OF PERSONS AND LINES AND A R. B.
                                                                                                                                                                                                                                                                                                                                                                                                                              and resident
26. 26. 25. 45. EM 1460. 1 March Construction Cons
  If $6.00 of $100 has been a superposed where the contract of t
  NAME OF TAXABLE PARTY OF TAXABLE ASSOCIATION
  A SERVICE FOR THE COME I CONSTRUCTION OF PROPERTY OF THE PROPERTY OF SERVICE STRUCTURE STRUCTURE
or Depart Addition for the Contract Assessment and Contract Age Contract Co
  regions, and consequently a state and have below, and despite falls, relegization, provides and
text and territories between the action and text and text and the contract of the text and the first and text a
  mai mula, decimento e Parlamentaria, presentaria televiseraria, formuniario e a. Parlam
THE REPORT OF THE PARTY AND A DESCRIPTION OF THE PARTY AND ADDRESS OF THE PARTY ADDRESS OF THE PARTY ADDRESS OF THE PARTY AND ADDRESS OF THE PARTY AND ADDRESS OF THE PARTY
     with a season to the contract of the contract 
OF SECURE AND ARREST ASSESS OF ADMINISTRAL PROPERTY AND ADMINISTRATION ADMINISTRAL PROPERTY AND ADMINISTRAL PROPERTY ADMINISTRAL PROPERTY AND ADMINISTRAL PROPERTY ADMINISTRATION ADMINISTRATION ADMINISTRATION ADMINISTRATION ADMINISTRATION ADMINISTRATION ADMINISTRATION ADMINISTRAL PROPERTY ADMINISTRATION ADMINISTRATION ADMINISTRAL PROPERTY 
married the eight a married parameter
  to be about your less of the control of the last and the 
  20 St. St. City City Cold communications among Sciences when University Chief St.
HE SELECT AN EVEN THESE I MANUFACTURE WHITE STATE I WINDOW AND A MANUFACTURE AND ADMINISTRAL A
I STANDARD LEAVE UND BEAUGUAGE MUSIC PRINTS AND A SECURIOR WAY BY
in the string are other critical in designationary function beautiful and streets for the second fraction of the contract of the second streets of the sec
  deat per proposate burner (Dell' del Abres) probe "rede la besta petral per moner basellera.
                                                                                                                                                                                                                                                                                                                                                                                                                              a shark that it has be completely and it has a
  are per decrease when he was the second
                                                                                                                                                                                                                                                                                                                                                                                                                                 E other than I bestuly business seems best-powered; hop-power
  to be deleted after 1 february 1 and 1 february 1 at 1 and 1
                                                                                                                                                                                                                                                                                                                                                                                                                              # (SR) NA constitut in Albert & wageting 1, interesting
  OF REPORTED AND THESE TRANSPORTS IN A SERVICE TO THE REPORT OF THE PARTY.
  IN DR. SCOTE PRO- DRIVE TODAY IN NAMED . . . IN CONTROL OF FRAME STATES AND AND
     NO DECEMBER OF THE REAL PROPERTY AND ADDRESS OF THE PROPERTY O
                                                                                     Witnesday but the season that being tracket
  to provide day. They set a foregraphy to be a few to b
  NATIONAL REPORT OF THE PARTY AND ADDRESS OF THE PARTY.
  to the first New York of the State of General Advantages of State 
  of the series with their areas or opening becomes their published in
```

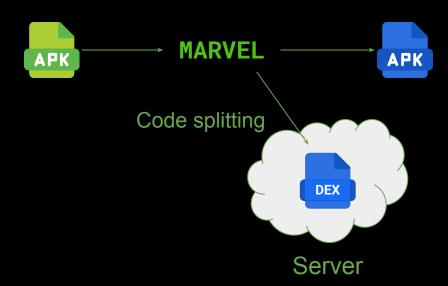


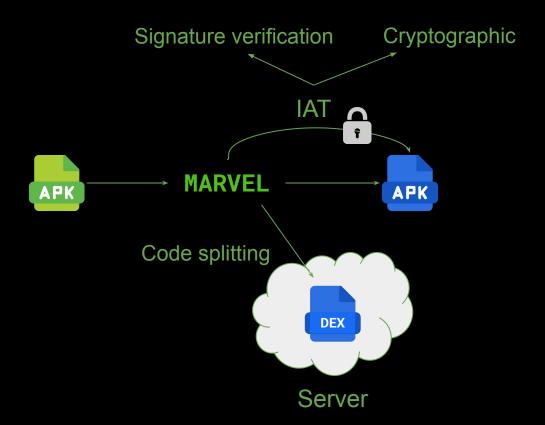
 We used Marvel to protect the 10 lightest apps among the top 5,000 most downloaded in the Google Play store

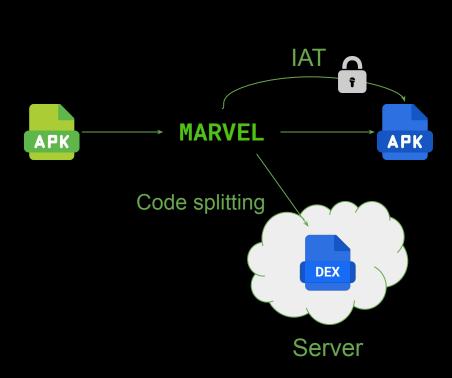




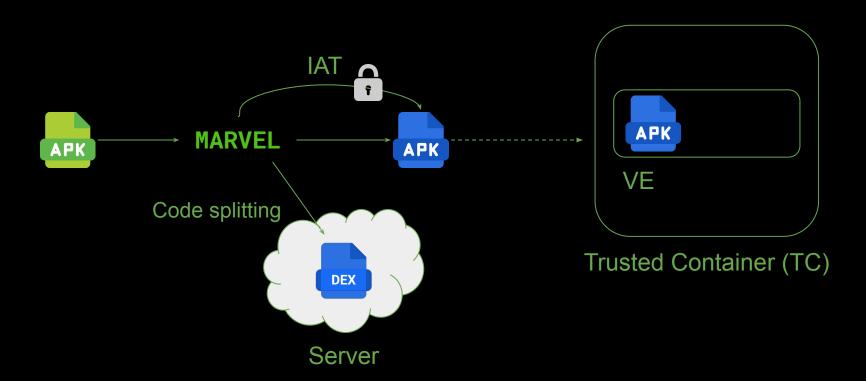


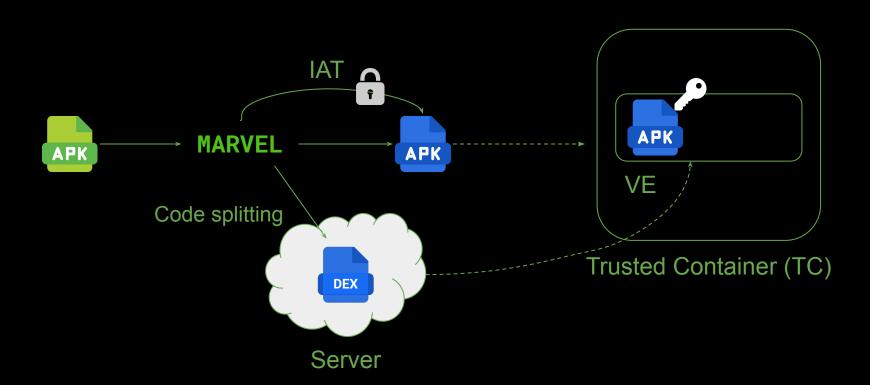


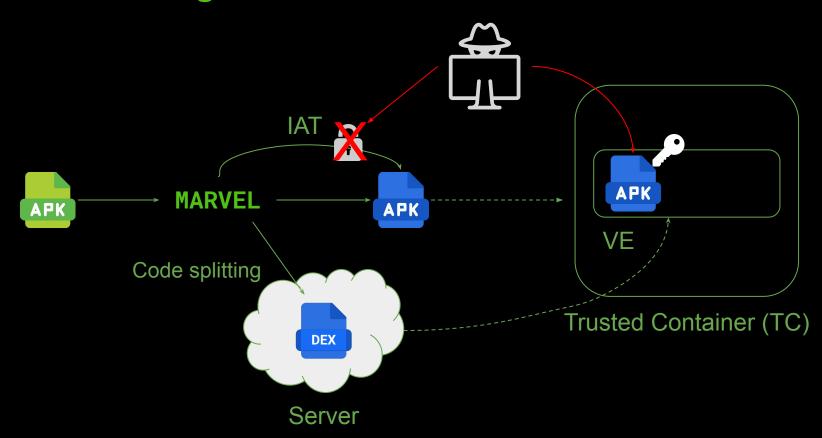


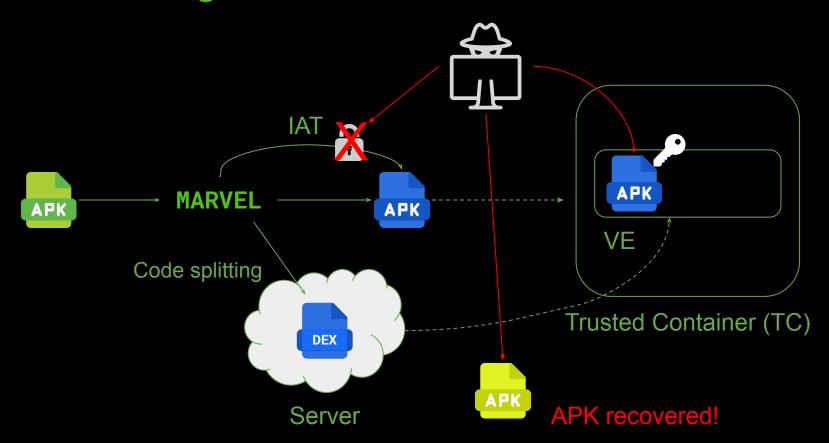














 Using Frida [7], we were able to intercept the code loaded and reconstruct the original applications

[7] https://frida.re/

- The TC copies the DEX files with the writeFileOnInternalStorage API
 - We hooked it to intercept them and locally dump them
- For IAT, we simply deleted the local class that performs the checks;)

We were able to **bypass** Marvel for **all of them** and allow execution outside the trusted container

RQ2: Defences adoption

- We executed the top 5,000 most downloaded apps in "2Account" [8], one of the most popular virtualization apps
- We collected the logs of crashing apps and reverse engineered apps that failed to execute

RQ2: Defences adoption

- We executed the top 5,000 most downloaded apps in "2Account" [8], one of the most popular virtualization apps
- We collected the logs of crashing apps and reverse engineered apps that failed to execute

None of them embedded any anti-virtualization measures or used Marvel

RQ2: Defences adoption

- 38 apps crashed in "2Accounts" for the following reasons:
 - Native library incompatibilities
 - Missing graphical resources
 - Incorrect file paths
 - 2Accounts bugs
- But no signs of virtualization detections

Matrioska's Motivation

We found several limitation in all existing defences

- The anti-virtualization libraries and Marvel are absent among the top 5k most downloaded apps
 - They require high involvement from developers
 - VAHunt is a **static** solution and cannot be directly deployed on user's **devices**

Matrioska's Design

We propose Matrioska, a dynamic defence tool that:

- Detects virtualization usage and repackaging attacks
- Avoids malware evasion and anti-debugging techniques
- Requires no modifications to the Android OS or root privileges
- Can be installed as a third-party app

Matrioska's Design

- Matrioska dynamically inspects an application against six signatures, then computes a score
- A high-score indicates the presence of virtualization and/or repackaging behaviour

Matrioska's Signatures

The first two are checked statically:

- S1: An APK is found in the "assets" directory
- S2: One or more "stub components" can be found in the app
 - Stub components serve as place holders for the plugins' components

Matrioska's Signatures

The rest are measured during execution:

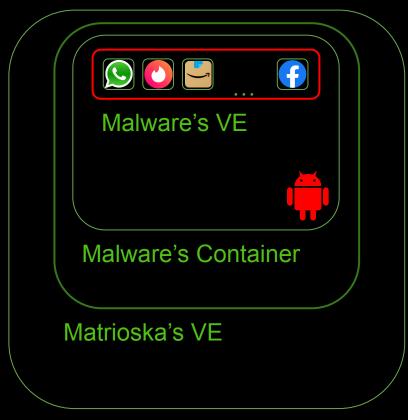
- **S3**: The app starts another app that is not installed in the devices (i.e., a plugin)
- S4: The app spawn processes sharing the same UID
- S5: An APK is downloaded or decrypted and saved in the storage
- **S6**: Container's signature differs from the plugin's

Matrioska's Design



Malware's Container

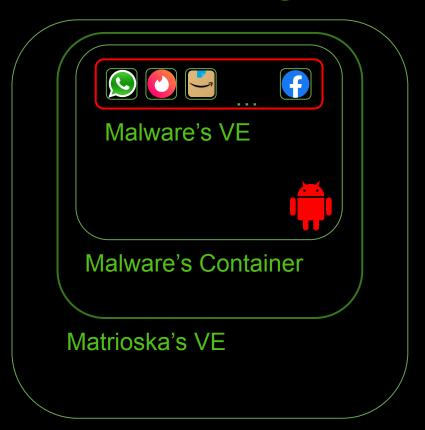
Matrioska's Design



Matrioska

42/58

Matrioska's Design





Source: Microsoft Copilot

Matrioska

- How well does our solution perform in detecting VBR malwares with respect to the state-of-art?
- We compared Matrioska against VAHunt, the state-of-the-art in detecting VBR malware

VAHunt

- VAHunt is a static analysis tool to detect VBR malware
- 1. Decompiles the APK into small code
- 2. Detects **virtualization engine** presence with hardcoded strings
- 3. Analyzes behavior for signs of **silent loading** a. API calls, call chains
- 4. Checks the Manifest, APK files in the "assets" directory

We used two different datasets:

 Dataset 1: 100 applications from the Google Play Store (50 with and 50 without virtualization) to measure virtualization detection

Virtualization Detection

Dataset 1	TP	TN	FP	FN	Accuracy
Matrioska	50	49	1	0	0.99
VAHunt	21	50	0	29	0.71

VAHunt fails to generalize because of the several hard-coded checks in its logic

- Dataset 2: VAHunt's authors refused to shared their dataset
- We collected 152,602 samples from AndroZoo [9] having a VirusTotal score equal to or greater than 10/66

 From 152,602 samples, we found a total of 1765 apps using virtualization

 We found 187 possible VBR malware, which we manually reverse engineered

VBR malware	FP	FN
Matrioska	10	14
VAHunt	23	39

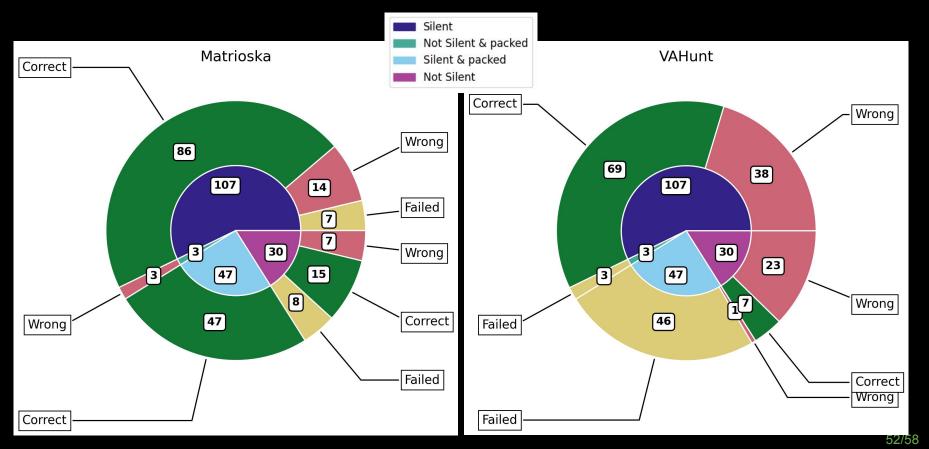
- Matrioska is still subject to some false positives and false negatives
- Some legitimate apps might have APKs in their assets directory, or use stub components
- Malicious apps could delay detonation

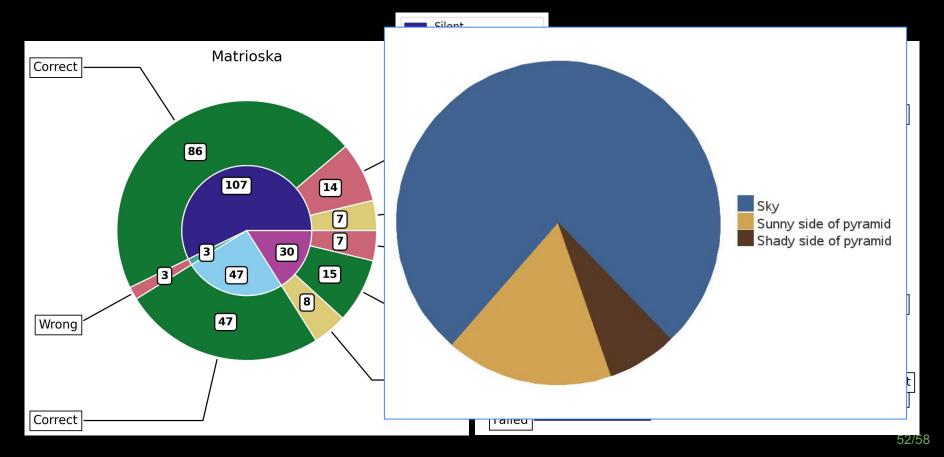
Video Demo

```
Claim Strike (S. Rocket T.C. SEC 1854 C. S. SENSON S. D. SERNAND ST. TOWN
    THE RESIDENCE OF STREET, STREE
    A DOMESTIC AND ADDRESS OF THE RESIDENCE AND ADDRESS OF THE PARTY OF TH
    A THE ROY OF SECURITIONS AND ADDRESS OF THE PARTY OF THE 
    COLD IN THE PERSON FOR THE REST LETTER SCHOOL OF REPORT AND ADDRESS OF THE PERSON OF T
    COLD STREET BEING FORD LITTLE STREET, BOOKS, STREET
    A DESCRIPTION OF SECURITIONS ASSESSED TO A DESCRIPTION OF THE PROPERTY AND A DESCRIPTION OF THE PROPERTY AND ADDRESS OF THE PROPERTY ADDRESS OF THE PROPERTY AND ADDRESS OF THE PROPERTY ADDRESS OF TH
    LOW TWO US TO BE IN THE THE REPORT OF THE PROPERTY OF THE PROP
    DO NOT READ THE ADD AND ADDRESS OF THE PARTY.
    by with both localist to bridge of
    DO DOWNERS IN AGRESSIONS
    THE RESERVE AND ADDRESS OF
(A.C.) ANALYSIS DESIGNATION OF SERVICE STANDARD CONTRACTOR STANDARD STANDAR
        PRODUCES CONTROL OF THE PRODUCE OF T
    removement introduced mornishing
    Markety: When I begrow, make a membership to be an a set of

    to reconstruction of the desirement of the property of the proper
    THE R WILLIAM ST. CO. SEC. SEC LINES LABOUR A DESCRIPTION OF THE PARTY OF
    Links New on Service for the complete Land of complete and conditional and of
    Link Name of the co. In the Latter Links of Helicarday and Artificial States
                                                                                            NAME OF REST OF STREET ASSESS OF CONCENTRAL ASSESSMENT 
Later A Value of the fire pr. Say Great Laters in contributions. Spendight Act.
Line Kirk of the children dead Least a minimum a problem. For
Later A year los per est as for Carbon Laters of Colonians in Colonian Confederation Colonians
Liver NAME 25 AN ET DE SON COME LANGE & COMMISSION OF COMMISSION OF THE PARTY OF TH
    THE WALL STREET, STREE
    THE WHEN SHE SHEETER BY STREETER
    CALL STREET, STATE BY STREET, STATE OF STATE OF STREET, STATE OF STATE OF STREET, STATE OF STATE OF STREET, STATE OF STATE OF STREET, STATE OF STATE OF STATE OF STREET, STATE OF STATE OF
        THY RESIDENCE
    CONTRACTOR STREET, AND ADDRESS OF THE PARTY 
            recentural transmission, introduction post facilities protected in this part of the process to a labor
        Tilland emerical projections
        August and
        CASTING THESE I DISCOULDED IN ANY TO ADMINISTRATE IN CONTRACTOR OF THE PARTY OF THE
            If was extreminate them in our descriptions with a set is contain blues the city
```







- VAHunt is limited by traditional static analysis limitations
 - It cannot analyze packed apps (i.e., apps decrypted at runtime)

Matrioska achieves **better performances** thanks to its dynamic analysis and signatures

Analyzing packed apps

- Packed APKs contain encrypted payloads or obfuscated code
- Medusa [10] is an open-source framework that can be used for dynamic DEX extraction
- Hooks into the ClassLoader at runtime to dump:
 - Decrypted DEX files
 - Loaded classes
 - de-obfuscated code before execution

 Finally, we validated a total of 154 VBR malware, in contrast to the 71,303 ones detected by the VAHunt's paper

VBR malware are an existing **threat**—maybe even a growing one—but are **not currently prevalent** at the scale prior work suggested

Limitations

- Code coverage
 - Matrioska assumes immediate malware execution post-startup
- Assumes plugins are silently loaded
 - Interactions with the UI may be required

Future works

- Increase code coverage
 - Localize payload and try to reach it
 - Employ LLM-guided UI interaction tools
- Develop more stable frameworks for virtualization and hooking
 - Some of these have very unstable development cycles

Conclusions

- Existing defences against VBR malware exhibit several limitations
- Matrioska achieves better performances against the state-of-the-art and can effectively safeguard Android users

Questions?



Source: Chat-GPT



Full paper