The Carbonara Project

Malware research platform and community

WHITEPAPER

Version 0.1

17th March 2018

Contents

1 Introduction

Malware research and reverse engineering are hard tasks.
They require a huge amount of technical knowledge, experience and creativity to be effective.

The amount of malicious programs, moreover, is increasing over time. Malwares become cheaper, and the attack surface becomes bigger and bigger.
In Q3 2016 alone, 18 million new malware samples were captured [1] .
Also, ransomware attacks increased by 36 percent in 2017 [2] and, according to the FBI [3], more than 4,000 ransomware attacks occur every day.

If the main targets of attacks in the past were personal computers and servers, a very high percentage of attacks is now targeting smartphones and IoT devices.

As the number of attacks increases over time, it looks like the number of people working against cyber crimes is not increasing proportionally.
Unfilled cyber security jobs are expected to reach 3.5 million by 2021 [4] — compared to about 1 million in 2016.

It is also important to understand that the job of most malware analysts is highly time-sensitive.

From the moment an attack is discovered, it is a race against time to understand how the malware works, find a fix, and apply it as soon as possible to as many vulnerable devices as possible.

We have developed a platform to speed up the work of reverse engineers and malware analysts. Our main goal is to create an environment in which security experts can thrive, and to empower them with state-of-the-art tools to make their job easier.

[1] Source: https://www.pandasecurity.com/mediacenter/pandalabs/pandalabs-q3/
[2] Source: https://www.symantec.com/security-center/threat-report
[3] Source:https://www.fbi.gov/file-repository/ransomware-prevention-and-response-for-cisos.pdf/view
[4] Source: https://cybersecurityventures.com/jobs/

2 The project

Carbonara is, at its core, a platform in which malware analysts can share their work, collaborate, and perform analysis using the tools we provide.
The users can load programs and malwares on our server and our engine will run a certain number of tools to perform static analysis.

Users can add their reports to binaries and to their procedures, as well as comment on them and discuss with other users.

One of the main features of our analysis engine is **procedure matching**.

What it does is matching similar procedures from different binaries. This is extremely useful: it has in fact being shown multiple times that a lot of different malwares of all categories share a lot of the same functions.

What this means is that, if a certain user is reversing a certain function, and that function has already been reversed by another analyst in some other binary, our engine will match the function, giving easy access to the report and explanation for that procedure.

We will also provide state-of-the-art machine learning algorithms to perform malware classification.

Carbonara is usable both from the web application and from the CLI (*Command line interface*).

3 Technical overview

A more technical explanation of our procedure matching system is available at *https://github.com/Carbonara-Project/Guanciale/wiki/The-matching-system-explained.*

Of course, for each binary we also produce the list of functions and the assembly code for each one on them, getting information from important disassemblers and static analysers such as **radare2** and **IDA Pro**.

The machine learning model will instead use a classifier in order to predict whether the loaded binaries are malwares and, if so, what type of malwares they are.

4 The community

The community is the main focus of our vision.
We are deeply convinced that there is not enough sharing of knowledge and information in the cyber-security community, which means that we are not as effective as we could be in stopping and fighting threats and attacks.

As we've stated in the introduction, the number of unfilled security jobs is huge and, with about 230,000 new malware samples produced every day [5] , it's very hard to keep up.

Among 50 security experts we surveyed, about 90% of them highlighted the lack of a community for the exchange of knowledge and solutions. Most of them also stated that their job is extremely time-sensitive.

This is why we want to make it extremely easier to share technical information on malwares, and to avoid repeating the work that someone else has already done in finding out how things work.

5 What's ahead

5.1 The Community

Our plan is to add several features to the interaction between users.
We need to introduce quality measures for the reports and anlyses made by the users, and we want to allow users to keep in touch with colleagues inside the platform.

5.2 The Analysis Engine

One of the main features we want to introduce is the ability to find more general similarities between entire malwares, instead of only comparing single procedures.

This would be highly beneficial, as analysts would benefit to great extent of work that has already been done.

We also want to make it easier to identify common functions in stripped bina-

[5] Source: https://www.pandasecurity.com/mediacenter/press-releases/all-recorded-malware-appeared-in-2015/

ries (which are much harder to work with), by creating plugins that work in conjunction with the most popular analysis tools.

6 The Team

Our team is entirely composed of computer engineering students from La Sapienza University of Rome.

We are also part of a CTF team called the TheRomanXpl0it (*https://theromanxpl0it.github.io/*).

This is us:

- **Daniele Cappuccio** - front-end developer
- **Luca Ferrera** - back-end developer
- **Andrea Fioraldi** - software engineer
- **Daniele Paliotta** - software engineer
- **Luigi Paolo Pileggi** - software engineer
- **Andrea Tulimiero** - back-end developer

7 Acknowledgements

We would like to thank our friends from *TheRomanXpl0it* for the support and for the knowledge they shared with us.

We also want to thank *Camil Demetrescu*, *Daniele Cono D'Elia* and *Emilio Coppa* for their teachings, and for providing us with a stimulating environment in which we could share our passions and ideas.