[27] Cybercrime Specialization: An Exposé of a Malicious Android Obfuscation-as-a-Service

Vit Sembera (Trend Micro), Masarah Paquet-Clouston (GoSecure), Sebastian Garcia (Czech Technical University) and Maria-Jose Erquiaga (Cisco).

Reviews Phase 1

Review 1

Time: Jun 07, 21:24

Overall evaluation: 2 (accept)

[Summary]

In this paper the authors investigated the workings of an obfuscation service geared towards mobile malware. The authors obtained access to this service, got a few known APK malware obfuscated, and then studied how the obfuscated malware was changed by the service. In this process, the authors discovered some mistakes done by that service that allowed them to find other malware obfuscated by the same service on VirusTotal. Using this method, the authors were able to estimate the revenue of the obfuscation service and characterize the actors that use it.

[Comments]

This is an interesting paper that characterizes the obfuscation behaviors of a specific service tailored for malware. I liked not only the breakdown of the offered functionality but also the fact that the authors were able to uncover issues with the service which they then took advantage in order to estimate the revenue of that service.

Some comments and thoughts for improving the paper:

- Why did you study that specific service? It would have been helpful to understand how that service compares to others in terms of popularity and cost.

- I would encourage the authors to incorporate network forensics in their analyses. E.g. how old was the domain hosting the service? Are there any other websites co-located on the same web server? Do the operators of that service need to move it from domain to domain because of take-downs, or is their login "wall" sufficient to evade detection and stay online for long period of time?

- Were there any differences in terms of sandbox evasions? I.e. does the website merely obfuscate the code for defeating static analyses tools, or do they also add anti-sandbox functionality to the uploaded malware?

- Since we know that VirusTotal shares samples with AV companies, it seems surprising that the service uses them, as opposed to VT clones that do not the same scans, but don't share executables. Do you have any theories about the use of VT in this case?

Reviewer's confidence: 4 (high)

Review 2

Time: Jun 21, 09:25

Overall evaluation: 1 (weak accept) *[SCORE CHANGED FROM 0 (borderline) AFTER REVIEWING THE REBUTTAL LETTER]*

> This paper presents a use case that looks into the technical details and revenue offered from a specific obfuscation service that is currently operating. The paper is generally well organised and covers a very important and timely topic. The methodology is reasonable but I found it hard to understand how the other obfuscated applications could be found in the wild. Is there any reason why 'radio.ogg' files could be potentially obfuscated? is this knowledge that was obtained during the analysis? Why did you focus on those ones? Also the existence of tracks seems to be a sign of the file but it is not clear to me how the authors arrive to that conclusion. Was this present in all the apps? I think more detail on how the authors determine the features of this obfuscation service are needed. These aspects are described in the methodology section but they seem to belong more to the result section right?
>
> I also found some of the details about the revenue lacking. I am not sure about the assumptions about the group of outliers. Are you considering they all purchased unique APKs? I think more detail could be provided here to help the reader understand how do you come up with those numbers.
>
> A minor issue regarding the vocabulary that should probably be cleared at the beginning. I found the usage of 'client' in this context confusing (maybe customer would be better?).
>
> However, my main concern has to do with the ethical aspects of the research. If the authors have gone through a research ethics committee (which I believe is the case) they should make this clear in the paper along with the outcome or comments the committee had on this approach.

Reviewer's confidence: 4 (high)

Rebuttal Letter

The text below is short due to the 600-word limit of the whole document. Do not hesitate to contact us if further clarification is needed.

1. Why specific service? (...)

This service was found while investigating attackers involved in a botnet. During the investigation a colleague found a chat log in VirusTotal, submitted by others, containing discussions among the attackers. In this chat, we discovered that the attackers used the service to obfuscate their files.

Regarding the comment on comparison, Section 6.4 does such a comparison and Table 4 compares the service with the top 6 competitors found in underground forums.

We will add information to the Introduction on how we encountered the service.

2. Network forensics (...)

There were two reasons why the forensic analysis was not included in the paper. First, because we wanted to focus on a business analysis and assess the impact of such a service. Second, we conducted a forensic analysis, but the results were not significant nor added much information to the analysis of the service.

We will add further information in Section 2 about the details of the IP address and its forensic analysis.

3. Sandbox evasions (...)

As far as we saw, sandbox evasion was not part of the obfuscation algorithm. The main purpose of the obfuscation service was to evade static AV engine detection.

We will update the description of the obfuscation to mention that no sandbox evasion was found.

4. Use of VirusTotal (...)

Those behind the obfuscation service did not use VirusTotal. Our paper states that VirusTotal was used by us to see if the obfuscated samples were detected. We also searched older obfuscated APK samples on VirusTotal and the APKs found were uploaded by companies, detection services and users.

We will add this clarification to the paper so there are no confusions.

5. Methodology(...)

Our analysis of the obfuscated APKs revealed that **inside the obfuscated APK** were parts that were in plain text, in particular the string of the filename 'radio.ogg'. These plain text parts were exactly *the same* for all the obfuscated APKs. We concluded that this string 'radio.ogg' was a very good 'Indicator of Compromise' and that it could be used in a YARA rule. Our methodology was to use this string in a YARA rule inside the VirusTotal Retrohunt service to find other APKs in the wild that had a string filename 'radio.ogg'.

We will add further description in the methodology and result section to clarify this situation and be clear that the string was used as an IoC in a VirusTotal search to find APKs.

6. Details Revenue lacking (...)

Yes, all outliers APKs are considered as purchased individually. Since there was no way for us to group them together, we hypothesized that they were probably generated by separate individuals and then used the price of one APK as reference.

We will further clarify in the methodology the rationale behind this decision, and our assumptions, together with the logical steps for our conclusion.

7. Vocabulary client (...)

We will make sure that the term is better defined throughout the paper.

8. Ethical aspect (...)

We agree with the reviewer that this research was delicate and needed authorization. We got the corresponding authorization from the Ethics Committee in our University in 2020, but we didn't send it before because of the anonymous submission. We sent the documents to the PC chairs.

We will modify Section 4.3 to improve the explanation as well as add a statement that ethics approval as granted.

Reviews Phase 2

Review 3

Time: Jun 27, 10:55

Overall evaluation: -1 (weak reject)

I enjoyed reading this paper. It explores an interesting topic and provides insights on a business that, even if it is relatively well known, it evolves and adapts regularly.

I am not more positive about this paper because of a number of methodological issues that makes me question some of the findings. Specifically:

- The number of samples is very low. While I understand the reason given by the authors (the service is expensive), drawing conclusions from just 3 samples is methodologically unsound. The whole analysis is based on gathering a dataset form VT based on signals obtained from the obfuscated samples (4.2.1) to assess the service usage (e.g., the radio.ogg file). I presume you identify those signals through reverse engineering of the obfuscated apps. (This is clear later after reading Section 5, so it would be good to describe it before). How can you be sure that _only_ the apps obfuscated by that particular service contain those signals? You claim that this is the case ``with a high degree of confidence,'' but this is unclear to me. You seem to be assuming that the service operators use a custom obfuscator that nobody else uses in the wild and that can be fingerprinted. What makes you think that this is indeed the case? A larger number of samples sent to the service over time would provide you with higher confidence.

- In terms of the samples chosen for the test, a sample cannot be considered adware just because it uses Flurry's SDK. There are legitimate uses for a monetization and advertising library such as that. Do you have evidence that the sample is truly adware?

- The grouping is also confusing. Why do you believe that apps with similar strings belong to the same service customer? You need to make this intuition explicit to better understand what your assumption is. When reading the results later in Section 6, you describe that apps within each groups behave similar (e.g., same network sinks). Did you did that automatically or is it based on visual inspection? Also, after reading the whole paper, I suspect the key reason for grouping the samples is estimating the number of customers for your economic analysis. It would be clearer if you state this upfront. Because of this, the estimation of the potential revenue is somewhat weak. I don't believe you can extrapolate how much money the service is doing from the observations in VT.

Another substantial comment has to do with ethical issues of the conducted research. Spending some money in underground criminal services is an accepted practice in the community if that contributes to a better understanding and the overall risk-benefit analysis is okay. In any case, if you are in an academic institution, you need to apply to your IRB. This is unclear in the paper and needs to be discussed _explicitly_.

Suggestion: I believe the paper presentation would improve if you merge sections 4 and 6. Having Section 5 in between breaks the reading flow. Section 5 contains very interesting insights on the operation of the obfuscator, but it's a bit disconnected from the analysis.

Reviewer's confidence: 4 (high)

Review 4

Time: Jun 30, 13:39

Overall evaluation: 2 (accept)

I have limited expertise regarding this article, as the methods are technical, which is not my background. This means that my comments relate to the framing and contribution of the paper. I am unable to offer meaningful points on the methods and results.

With my limited expertise, overall this seemed a solid paper to me. Though I do think some improvements could be made. Most importantly, the focus and contribution could be clearer. The introduction seems to overplay the novelty of the article in relation to specialization ("the mechanics behind the scene have yet to be studied"), although there is a greater acknowledgment of existing literature on specialisation in following paragraphs. The paper is missing some well-known studies of cybercriminal specialisation, e.g. Grier et al., 2012; Lusthaus, 2018, chapter 3; Manky, 2013; Collier et al., 2021 (already cited, but not directly on specialisation). There are obviously very many law enforcement and industry reports that address the related topic of cybercrime as a service. It would be good to make narrower statements upfront about potential contribution, to discuss the existing literature on specialisation in more detail and better explain what is known about cybercriminal specialisation already, and what still needs to be uncovered. Does a study of obfuscation specifically address these gaps? Or is such a study part of a broader work program that could analyse any number of topics with the specialisation framework, and obfuscation is simply one of the areas that has not been studied? Apart from such a gap, are there more specific reasons to study obfuscation, as opposed to any other topics which have not had the specialisation framework applied to them?

I found the opening sections to be oddly structured. I don't think it is needed to list the key takeaways in the introduction. This is best left to the conclusion (and abstract if desired). It also feels like there are too many opening sections before we get to the methods. I wonder if better focus could be achieved by simply having one literature review section following the introduction, which more clearly outlines the motivation of the study and its intended contribution within the field.

Tied to these framing points, in the discussion and conclusion, I wanted to see a stronger link to what the results teach us about cybercriminal specialisation overall. What do we know now, that we didn't know before this analysis? It would be good to be a little more analytical here, and engage more with how this work confirms, contradicts and/or extends the existing literature. At the moment, these elements are largely descriptive. It sometimes feels as if certain elements are repeated multiple times, but what is less clear is why these outcomes matter to the field. This should be made clearer.

A secondary point (and this should remain secondary, as the focus needs to be much tighter as to what this article primarily contributes on specialisation) is around the discussion of profits. This is an area that many scholars have struggled to address. So the fact the authors have something to say on this point should be highlighted and better connected to some past attempts to address this question, and the challenges therein (e.g. Holt et al., 2016; Dupont & Lusthaus, 2021). Does the authors' contribution add more certainty on this point regarding cybercriminal profits?

Reviewer's confidence: 2 (low)

Review 5

Time: Jul 02, 10:54

Overall evaluation: 2 (accept)

=== meta review ===

The PC agrees that this paper makes for an interesting contribution to the workshop and should be accepted. However the reviewers also provide a number of suggestions and clarifications to improve the paper. In particular, please provide explicit details about the ethical issues, including the involvement with your institutional Ethics Board.

Reviewer's confidence: 5 (expert)