Human Factors in Cybersecurity, Vol. 53, 2022, 44–48
https://doi.org/10.54941/ahfe1002200

Isolating Key Phrases to Identify Ransomware Attackers

Jeremy Blackstone[1] and Wayne Patterson[2]

[1]Howard University, Washington, DC, USA
[2]Patterson and Associates, Washington, DC, USA

ABSTRACT

Ransomware attacks are a devastatingly severe class of cyber-attacks capable of crippling an organization through disrupting operations or egregious financial demands. A number of solutions have been proposed to decrease the risk of ransomware infection or detect ransomware once a system has been infected. However, these proposed solutions do not address the root of the problem: identifying the adversary that created them. This study takes steps towards identifying an adversary by utilizing linguistic analysis of ransomware messages to ascertain the adversary's language of origin. Our proposed method begins by using existing ransomware messages. We isolate commonly used phrases by analyzing a number of notable ransomware attacks: CryptoLocker, Locky, Petya, Ryuk, WannaCry, Cerber, GandCrab, SamSam, Bad Rabbit, and TeslaCrypt. Afterwards, we translate these phrases from English to another language and then back to English using Google Translate and calculate the Levenshtein Distance between the two English phrases. Next, we identify the languages that have a Levenshtein Distance greater than 0 for these phrases due to differences in how parts of speech are implemented in the respective languages. Finally, we analyze new ransomware messages and rank the languages from easiest to most difficult to distinguish.

Keywords: Ransomware, Cyberattack, Language, Levenshtein distance

INTRODUCTION

Ransomware attacks are a class of cyberattacks that prevent a user from accessing data on their device until the user provides some type of compensation. There are millions of cases of ransomware every year and advancing this field will protect consumers, private companies and government organizations (Simoiu, 2019). In this attack, the adversary provides the user with notification of the attack, what is necessary to recover from the attack and the steps they must follow in order to restore the original state of their device. However, it is possible for the adversary and the target user to have a different language of origin than the language used in the ransom message. In light of this, the purpose of this paper is to provide a methodology to determine the language of origin for a ransom message in a ransomware attack. While this approach has been proposed in (Patterson, 2022), this study focuses on text from notable ransomware messages and uses a wider variety of languages.

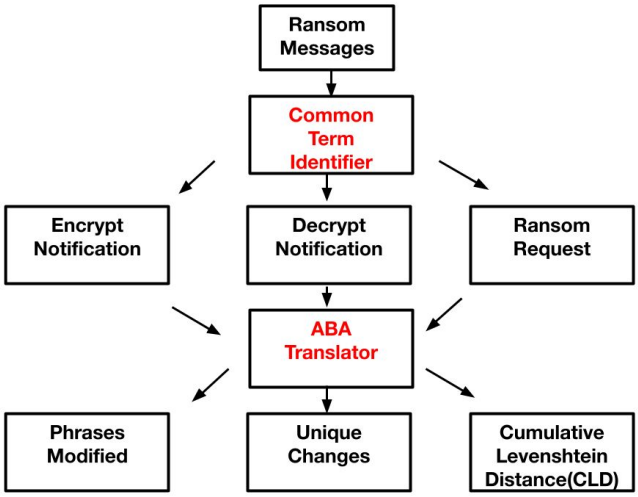

Figure 1: Workflow for isolating and analyzing key phrases.

We conduct our investigation based on the process shown in Figure 1. We begin by analyzing the text of ransom messages used for notable ransomware attacks and identifying common terms used in all of the messages. Next, we extrapolate these terms into three common phrases used in all ransom messages: the encryption notification, the decryption notification, and the ransom request. Afterwards we do a translation of this message from English to another language, then back to English (ABA translation). Once we have determined the degree of difference between phrases and why they are different, we use this information to rank how easy it is to distinguish the languages based on the phrases modified, the number of unique changes made and the cumulative Levenshtein Distance (CLD).

The major contributions of this paper are:

1) Identifying common phrases used among a variety of ransomware messages
2) Determining the degree of differences between foreign languages and English for common phrases used among a variety of ransomware messages
3) Generate a list of languages capable of being identified based on ransomware messages organized from easiest to most difficult to distinguish

BACKGROUND

Threat Model

We assume that the adversary is a non-native English speaker and that the final ransomware message is in English translated via Google Translate. We assume this because using a language pervasive throughout the internet provides an adversary with more potential targets and a large volume of ransom notes are translated via Google Translate as shown in (Florea and Patterson, 2021).

Levenshtein Distance and ABA Translation

Levenshtein Distance(LD) is a technique used in information theory used to quantify the difference between two string sequences (Levenshtein 1966). We propose using LD as a metric to quantify the degree of differences in phrases before and after translating them to another language. In this study we perform an ABA translation by translating an English message to another language, then back to English and calculate the LD as shown in example 1.

Example 1: English-Chinese(Simplified)-English
A: To decrypt your files
B: 解密你的文件
C : To decrypt [**your**] file[**s**]
LD: 'y'+'o'+'u'+'r'+' '+ 's' = 6

EVALUATION

Identifying Key Phrases

In order to isolate key phrases used in ransomware messages, we began by determining which terms were common among a sample of ransom messages from the notable ransomware attacks CryptoLocker, Locky, Petya, Ryuk, WannaCry, Cerber, GandCrab, SamSam, Bad Rabbit, and TeslaCrypt. To perform this analysis we stored the text of the ransom messages as text files and used a python script to read the files, store the words from each file as a list then determine the intersection of each pair of files to determine which words were most common. We found that almost every message used the words "encrypted" and "decrypt". We also found that half the messages have the word "pay". Given this discovery we decide to analyze the phrases within the messages where these words were used to generate a basic generalized ransom message as shown in Table 1. The three key words were connected to the 3 main components present in each ransom message: the encryption notification, the decryption notification and the ransom request. The encryption notification informs the user that they are no longer able to access their files ("Your files are encrypted"), the decryption notification asserts that there is only one option to regain access to their files("To decrypt your files") and the ransom request demands a price to enable that access ("You need to pay").

Analyzing Key Differences

After generalizing our generalized ransom message, we performed an ABA translation and calculated the LD [for each language. Finally, we organized the languages from most to least distinguishable] as shown in Table 2. In this analysis, we excluded cases where the only difference was replacing the words "encrypt", "decrypt" or "pay" with a synonym. We did this because this type of change does not alter the part of speech or structure of the phrase in any meaningful way.

First we note that the **Uyghur** was the most easily distinguishable language. It had the highest cumulative Levenshtein Distance(CLD) for the three phrases(17), changed both the encryption notification and the decryption

Table 1. Mapping common terms to phrases.

Term	Encrypt	Decrypt	Pay
Component Phrase	Encryption notification Your files are encrypted	Decryption notification To decrypt your files	Ransom request You need to pay

Table 2. Most easily distinguishable languages.

Language	# Phrases Modified	# Unique Changes	# Unique Combinations	Cumulative Levenshtein Distance (CLD)
uyghur	2	2	1	17
chinese	2	2	1	12
slovenian	2	1	1	16
hungarian	2	1	1	10
korean	2	0	1	9
kyrgyz	1	1	0	15
azerbaijani	1	1	0	11
maori	1	1	0	7
czech	1	1	0	7
hawaiian	1	1	0	4
arabic	1	1	0	2

notification phrases and had unique changes for both. We classified **Chinese** as the second most easily distinguishable language.. Although it did not have the next highest CLD(12), it changed both the encryption notification and the decryption notification phrases and had unique changes for both. We determined that **Slovenian** and **Hungarian** were the third and fourth most easily distinguishable languages because each of them changed two of the three phrases and had a unique change for at least one of them. Slovenian changed the decryption notification and ransom request and had a unique change for the ransom request and Hungarian changed the encryption notification and the decryption notification and had a unique change for the encryption notification. We chose **Korean** as the fifth most easily distinguishable. We did this because although it did not generate any unique changes, it changed both the encryption notification and the decryption notification with a unique combination of changes. It was the only language to change the plurality of the encryption notification and delete the word "your" from the decryption notification. Subsequently, we decided k**yrgyz, azerbaijani, maori, czech, hawaiian and arabic** as the next most easily distinguishable because they made a unique change to at least one of the three phrases.

We classified the remaining languages with CLD greater than 0 as the least distinguishable in our study. These languages made changes to the phrases, but none of the phrases were unique so a single language of origin could not be uniquely identified. However, it is possible to reduce the scope of potential adversaries to a particular region. First we found that primarily eastern european languages changed an infinitive verb phrase to an imperative verb phrase in the ransom request and deleted the word "your"in the

decryption notification. Second, we found that primarily southeast asian and pacific islander languages changed the plurality of the encryption notification and decryption notification. However, we found that Sudanese and Icelandic languages made these changes as well even though they are from very different geographic regions. Finally, we found that only pashto and punjabi deleted the word "are" from the encryption notification.

CONCLUSION

This study demonstrates how ransomware messages are composed of three essential elements: an encryption notification, a decryption notification and a ransom request. We generated phrases for each of these elements in English and performed an ABA translation on all languages included on Google Translate and found that select languages differ from the original message because the languages change parts of speech, add or delete words, change the verb tense, or change words from their plural to singular form. Using this information, we can identify the language of origin for a new message based on whether it prefers or omits the types of grammar we observed in our analysis. While this study focused on ransomware messages, similar analysis could be applied to phishing messages as indicated in (Patterson and Blackstone, 2022).

REFERENCES

Florea, D., Patterson, W. (2021). A Linguistic Analysis Metric in Detecting Ransomware Cyber-Attacks, www.thesai.org

Levenshtein, V.I., 1966, February. Binary codes capable of correcting deletions, insertions, and reversals. In Soviet physics doklady (Vol. 10, No. 8, pp. 707–710).

Patterson, W. (2022). Detecting Cyberattacks Using Linguistic Analysis. Proceedings of the 13th International Conference on Applied Human Factors and Ergonomics (AHFE 2022), New York, NY.

Patterson, W., Blackstone, J. (2022). A Metric To Assist in Detecting International Phishing or Ransomware Attacks. Proceedings of the 13th International Conference on Applied Human Factors and Ergonomics (AHFE 2022), New York, NY.

Simoiu, C., Bonneau, J., Gates, C.and Goel, S. I was told to buy a software or lose my computer. I ignored it: A study of ransomware. In Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019), pp. 155-174. 2019.