Detection of Malicious Network Traffic Behavior Using JA3 Fingerprints

Pavel Novák¹, and Václav Oujezský²

¹²Faculty of Informatics – Masaryk University, Czech Republic

E-mail: 468940@mail.muni.cz, oujezsky@fi.muni.cz

Abstract—This paper presents a novel approach for classifying spoof network traffic based on JA3 fingerprint clustering. In particular, it concerns the detection of so-called zero-day malware. The proposed method does not work with known JA3 hashes. However, it compares the JA3 fingerprint of captured traffic with JA3 fingerprints of traffic with predefined criteria, such as the use of current cipher suites or protocol, for classification.

Keywords—clustering, detection, JA3, JA3s, malware

1. INTRODUCTION

The amount of encrypted network traffic has increased dramatically in recent years. It is valuable to protect users’ data against eavesdropping or modification. Nevertheless, encryption also opened new opportunities for malware developers because it makes hiding malicious communication more effortless. This problem has become much more critical because the portion of CC (Command and Control) server encrypted communication generated by the malware was around 45% last year, and the number is increasing every year. Encrypted traffic is much more harder to analyze because of privacy issues and performance demands. This inevitably leads to massive research in this area and the development of techniques to classify encrypted traffic.

While encrypted traffic is beneficial for legitimate users and protects user data, it also makes it easier to hide traffic generated by malware. There can be several types of such malware traffic. It can be network activity caused by worms, phishing campaigns, or communication with command and control servers. Encrypting such traffic effectively eliminates detections based on known signatures, such as detecting specific text strings or patterns in the communication. An encrypted traffic inspection is usually tricky because of privacy concerns and the difficulty of the decryption process. A possible solution is behavioral methods. These analyze the behavior of the communicating parties and look for anomalies. The second group of methods is based on traffic fingerprinting or monitoring the basic parameters of communication that occur in unencrypted form. One particular method is the JA3 fingerprint [1]. JA3 collects parameters that occur in unencrypted form while establishing a secure connection. However, JA3-based detections are based on direct matches. Thus, malware is not detected in the case of an unknown fingerprint. For example, there is currently a collection of JA3 fingerprints that have been obtained by analyzing more than 25,000,000 PCAPs generated by malware samples [2]. As the source states, these samples have not been tested against “good” traffic, i.e., they do not ensure malware detection.

The novel method described in this paper attempts to address this problem. The core of this work is to verify the ability to detect unsolicited encrypted traffic based on the similarity of JA3 fingerprints and to analyze what differences exist during the establishment of a secure channel of benign and malignant parties. Therefore, the method proposed in this paper uses the selected clustering methods to analyze and cluster JA3 fingerprints. Results of these methods will be analyzed and compared. This research aims to experimentally verify whether this detection method is applicable and which algorithm is best suited for this task. Accuracy, performance, and usability on a large volume of data will be assessed.

2. JA3 METHOD AND JA3 OUTPUT FORMAT

JA3, the version of the JA3 for the server fingerprinting JA3s, and the active version of JA3 – JARM are three fingerprinting methods developed by John Althouse and his team in 2017 [1, 3]. The JA3 fingerprint is created from data obtained from the ClientHello Message when establishing a secured connection. The specific values extracted from this message are the TLS (Transport Layer Security) version, accepted ciphers, list of extensions, elliptic curves, and elliptic curve format. Values from the specified fields are represented as strings and then hashed using the MD5 (Message-Digest) hash function. The MD5 hash is also the resulting JA3 fingerprint.
The output of JA3 is a text string of variable length. It contains five sections separated by a hyphen. A comma separates the individual values. Due to the variable length, the MD5 hash of JA3 is usually used as the resulting fingerprint in practice. Figure 1 shows the schema of the JA3 fingerprinting process.

Due to the nature of the JA3 fingerprint as a text string, clustering methods and metrics suitable for this type of output were considered in the design of the method. A more specific description is given in Section 3.

3. APPLICABLE CLUSTERING ALGORITHMS

Clustering is the task of dividing input data into similar groups. The division is done so that the points assigned to the same group are “more similar” in some sense. Clustering can be divided into two basic groups. The first group of methods is called hard clustering methods. These methods assign each sample to one cluster, i.e., for each cluster and each sample, the given sample belongs to the cluster or not. The opposite approach is called soft clustering methods. These methods can assign one sample to multiple clusters with some probability [4]. An overview of clustering algorithms that could be used is below.

K-means method, e.g., Lloyd algorithm iteratively divides the set of samples into \( k \) disjoint clusters. Each cluster has a centroid \( \mu_j \), which is a mean of all samples within the cluster. In the first iteration, the centroids are usually randomly selected. In each iteration, clusters are created by assigning a sample to the closest centroid. The centroid position is then adjusted, so it is the mean of the entire cluster. This process is repeated until the centroid position is not changing significantly. This algorithm is not optimal because a result can be only the local optimum [5].

OPTICS (Ordering Points to Identify the Clustering Structure) is a density-based clustering algorithm. It is a generalization of another well-known clustering algorithm – DBSCAN (Density-based Spatial Clustering of Applications with Noise). This algorithm arranges the samples in order in such a way that the two closest samples are always immediately after each other. The resulting clusters are determined by splitting the sequence at points where the relative distance exceeds a given threshold. This approach has the advantage over the similar DBSCAN algorithm of identifying clusters with different densities [6, 7].

Brown clustering is usually applied when working with text and analyzing context. The method is based on the assumption that words of similar meaning occur in the text in a similar context. This means that the distribution of words that immediately precede or follow are similar for similar words. This method is agglomerative. In the beginning, each word is in its cluster, and gradually, the clusters merge [8].

In order to measure the distance between samples, three discrete metrics – the Levenshtein, Damerau-Levenshtein, and Hamming metrics, will be considered in the context of clustering. These metrics will be used in the normalized form as the length of the strings in the different parts of the JA3 fingerprints are different and could potentially bias the results [9].

4. EARLY RESULTS

The initial part of the research aimed to verify two hypotheses about malware secure communication. The first hypothesis was that encrypted malware communication is similar across families. The second hypothesis was that the JA3 fingerprint parameters of malware-generated traffic differ from the JA3 fingerprints of benign traffic. Initial validation was performed using two test datasets. Both datasets contained the same set of randomly selected JA3 fingerprints of benign traffic. These samples were obtained from the JA3 database [10]. The remaining JA3 fingerprints were obtained by using analysis
of pcap files from the malware database [11]. The first test dataset contained communication samples of different versions of Qakbot, Emotet, Dridex, and Cobalt Strike malware. The clustering algorithm used in this phase was OPTICS along with the Damerau-Levenshtein normalized metric applied on individual parts of JA3 separately and summed together. The OPTICS algorithm was chosen because it does not require any prior knowledge about clusters in the dataset, and it is easy to extract clustering information from the ordering. Figure 2 shows the reachability plot as a result of applying the OPTICS algorithm to the first dataset. Red points indicate malware samples, and grey points correspond to malignant communication.

This algorithm sorts samples so that the two closest samples are next to each other. The x-axis corresponds to the ordering of points, and the y-axis corresponds to the distance of the nearest sample. In other words, the malware samples form the densest group in the whole dataset. It confirms the first hypothesis, namely that malware communication within a family is very similar. The second observation is that the malware samples form a separate cluster, which is quite distant from neighboring clusters. It confirms the second hypothesis. Several malware samples were outside this cluster, but they were also quite different from the benign traffic. Thus, they can be considered as outliers.

The second dataset contained samples of RAT malware, specifically the NetWire and the Revenge RAT [11]. The results were similar to the first case. Malware samples formed a separate group with similar JA3 fingerprints, which were quite different from the benign traffic. From this perspective, the results look pretty promising in terms of possible applications in detection. Since the malware samples perform similarities in terms of their JA3 hashes, we would be able to use the K nearest neighbors classification algorithm to recognize unknown malicious traffic.

5. DISCUSSION

This method could significantly help detect zero-day malware and identify its traffic on the network in the future. Moreover, this method does not require any disclosure of the privacy of the communicating parties. Detection is based on freely observable attributes and does not disclose users’ privacy. This method is also relatively easy to implement. JA3 fingerprints can be collected by most network probes currently in use. Moreover, their collection and computation are neither computationally nor memory intensive. Moreover, JA3 fingerprints are currently implemented in several intrusion detection systems, like Suricata.

This method, of course, has its drawbacks as well. The most important one is its relative ease of deception. If the malware mimics the TLS handshake of a legitimate program, the communication will not be correctly detected by this method. Another possible disadvantage is the incorrect result of the used clustering algorithms. In particular, the K-means algorithm is very susceptible to influence by outliers. It can then lead to suboptimal clustering and incorrect results.
Several possible extensions and modifications of the proposed method can be suggested for future improvements by considering the mentioned disadvantages. One possibility is combining this method with other zero-day malware detection methods. Specifically, these include algorithms that look for anomalies in network behavior. The second option is to use another fingerprinting method that contains more information than JA3. One example of such could be the Mercury method [12]. In addition to the attributes present in JA3, Mercury also uses all the TLS handshake extensions. One of them is a Server Name Identifier, which provides the server name with which the client communicates. Using such information could help to detect malicious traffic more precisely. However, Mercury is not as widely used as JA3 nowadays.

6. CONCLUSION

This paper presented the proposal of a novel method for classifying encrypted traffic based on the similarity of JA3 fingerprints. This work aims to verify if there are differences in the TLS channel establishment process between malware and benign traffic that could be used for traffic classification. The research involves comparing several selected clustering algorithms. This paper presents the first results of experiments performed with the OPTICS clustering algorithm. The first results have shown differences in JA3 fingerprints compared to benign traffic, which could be used to classify the encrypted traffic.

REFERENCES