CoDoH : Classification of DNS over HTTPS

D Balajee, C Siddarth, A Navaas Roshan, B Bavesh
Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram

Abstract

With the growth of internet services, computer network attacks are also equally growing. HTTP/HTTPS is the protocol widely used for transferring client requests to the server. But this protocol is prone to cyber attacks which is an alarming threat to the users' privacy. DNS over HTTPS was recently introduced providing security to the client requests' by preventing middle-man attacks. All DoH traffics are not as genuine as they should be, thus we have figured a technique incorporating some of the features of the DoH traffic to distinguish benign vs malicious DoH.

1. Introduction

HyperText Transfer Protocol (HTTP) is a straightforward protocol to communicate the client's request to the server. A web server receives an HTTP request that is sent by the client. The server processes the request and returns the respective HTTP response which will be received by the client. Due to its simplicity, in this protocol attackers can eavesdrop on sensitive and highly confidential information like payment details, etc. This led to the introduction of HTTPS. The later protocol encrypts the client requests and sends them to the server, so the intruder may receive the data but would be not able to extract meaningful information out of it. DNS stands for Domain Name System which translates the domain names (like youtube.com, google.com, etc.,) to the respective IP addresses that would enable the browser to load the resources from that particular website.

During the process of resolving an IP address, it would be known to multiple DNS servers for what domain we are searching for. This proves that DNS can also be exploited by cybercriminals. This ultimately leads us to DNS over HTTPS (DoH). By combining the technique of encryp[tion we used in HTTPS with DNS will ensure that the request is not tampered with or eavesdropped on by the DNS resolver or on-path routers. A setback is that hackers could generate malicious DoH traffic, i.e. TCP traffic encapsulated DNS queries are sent.

We perform DoH network traffic classification using machine learning based on the statistical features of the request. The first part uses a Logistic Regression algorithm on the features of the request to distinguish if its DoH is. Further, at the second layer, if it is found to be DoH, a similar algorithm carries out the classification of benign vs malicious DoH.

2. Related Work

Mohammadreza et al. [1] has attempted to detect and characterize DoH traffic in an online environment. This was one of the first works on DoH traffic detection and characterization. Mohammadreza et al. created a dataset, named CIRA-CIC-DoHBrw-2020 consisting of HTTPS traffic flows with two levels of distinct labels. It has DoH and Non-DoH HTTPS traffic at the first layer and in the next layer, the DoH traffic is segregated as Benign DoH and Malicious DoH (second contribution).

Yaser M. Banadaki [2] presented a systematic two-layer approach for detecting DoH traffic and distinguishing Benign-DoH traffic from Malicious-DoH traffic using six machine learning algorithms.

We found a major flaw with the feature selection of the machine learning algorithms of the above two works, which will be rectified in this work. The work is structured as follows: the dataset is explained in section 3, followed by our approach in section 4. The results are detailed and justified in section 5, with the conclusion in section 6.

3. Dataset

The CIRA-CIC-DoHBrw-2020 dataset was created using a two-layered approach that captured both benign and malicious DoH traffic, as well as non-DoH traffic. HTTPS (benign DoH and non-DoH) and DoH traffic were generated by visiting the top 10,000 Alexa websites and using browsers and DNS tunneling tools that support the DoH protocol, respectively.

Non-DoH: Non-DoH traffic was captured and labeled as traffic generated by accessing websites that use the HTTPS protocol. Thousands of Alexa domain websites were browsed to capture enough traffic to balance the dataset.

Benign-DoH: Benign DoH traffic is non-malicious DoH traffic generated by using the same technique as non-DoH traffic, namely the Mozilla Firefox and Google Chrome web browsers.

Malicious-DoH: To generate malicious DoH traffic, DNS tunneling tools such as dns2tcp, DNSCat2, and Iodine were used. TCP traffic encapsulated in DNS queries was sent by these tools. As a result, DNS queries were sent to special DoH servers via TLS-encrypted HTTPS requests.

Class	Number of samples
Non-DoH	897493
DoH	269643
Total	1167136

Table 1 a. Class count for task 1

Class	Number of samples
Benign DoH	19807
Malicious DoH	249836
Total	269643

Table 1 b. Class count for task2

The statistics of the dataset are shown in table 1.

4. PROPOSED METHODOLOGY

A. Preprocessing

The dataset consists of 28 features that could be used to find the class of the given request. Out of the 28 features, 25 features are float and 3 were of non-integer type, namely SourceIP, DestinationIP, TimeStamp.
The SourceIP and DestinationIP are split into 4 parts based on the 'period' in them. The first 3 parts correspond to the Network ID, and the last part corresponds to the Host ID.
The timestamp present in the format of '%Y-%m-%d %H:%M:%S' was broken down into simpler parts such as year, month, etc.
Upon inspection, we found that there was a bias created between the date of the packet request and the request's actual class label.

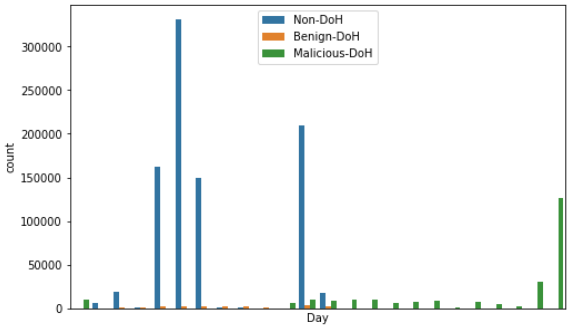

Figure 1: Bar plot of requests made on each day

Figure 1 is a bar plot of the counts of the requests captured on each day, hued by the request's nature (DoH vs Non-DoH). It is evident from figure 1 that the majority of the Non-DoH requests were placed in the first half of the month. Whereas the majority of the Non-DoH requests were in the second half of the month. This creates a bias in the model and may cause erroneous predictions in a different dataset. This was not considered in the previous work, which may have caused the model to bias its output on an irrelevant feature. Hence we drop all the features related to timestamp.

Highly correlated features were dropped with the threshold set as 0.75, except SourceIP1 and SourceIP4. Even though both are highly correlated, this is because all requests were made from the same subnet, but may come in handy when testing from an unseen sub-net. Finally, the features were normalized before passing into the machine learning model.

B. Model training

The available dataset was split into 90% for training and 10% for validation/testing.

Two separate Logistic Regressions were used to classify the input data. A logistic regression model predicts the dependent data variable (label) by analyzing the relationship between one or more existing independent variables (features). Logistic regression was our choice of model owing to its less computational complexity and efficiency.
Two different models were trained using the available data and a confusion matrix was constructed due to the natural data

	P	**R**	**F1**	**A**
Non-DoH	0.99	0.99	0.99	0.98
DoH	0.95	0.95	0.95	

2a : Non-DoH vs DoH

	P	R	F1	A
Benign	0.94	0.79	0.86	0.98
Malicious	0.98	1.00	0.99	

2b : Benign vs Malicious DoH

Table 2 : Confusion Matrix

imbalance in the data set. The classification report is discussed in the next section.

5. Results

It is evident from Table 2 that Logistic regression handles the classification very well. An F1 score of (0.99, 0.95) was achieved in classifying requests between DoH and Non-DoH. The second model achieves an F1 score of (0.86, 0.99). Both the models acquire an accuracy of .98, which is very high but the F1 score would be a better metric.

Figure 4 is a set of four figures consisting of influential features which affect a data point's class.

Figure 4a corresponds to the top 5 features that determine the input to be a Non-DoH sample. The traffic generated by accessing a website that uses HTTPS protocol is captured and labeled as non-DoH traffic. These are TTPS requests done in port 443, unlike DNS requests HTTPS requests vary depending on the user, hence it depends more on the FlowBytesRecieved and PacketLength.

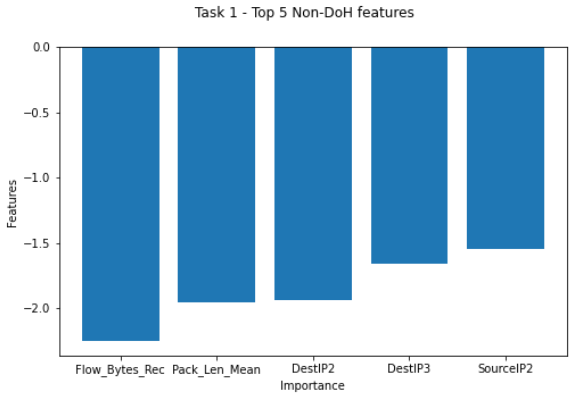

Fig 4a: Top 5 distinguishing features for Non-DoH

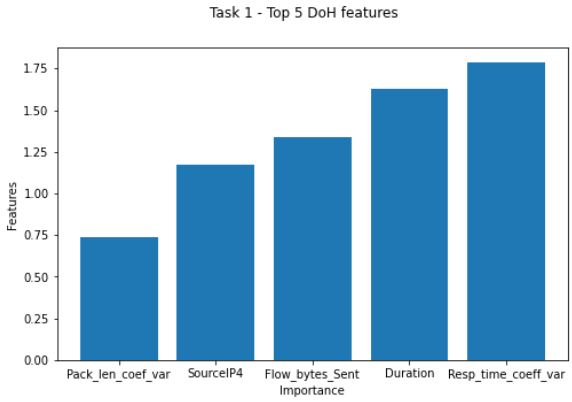

Fig 4b: Top 5 distinguishing features for DoH

Figure 4b corresponds to the top 5 features that determine the input to be a DoH request. The influence of the response time coefficient is justified as DoH operates over TCP, which can re-transmit data very quickly in the case of packet losses, whereas traditional DNS clients use UDP and wait for a fixed time before retrying. So in lossy networks, DoH may outperform UDP-based DNS. Hence the dependency is more on the RTCV and duration. Added to this the lesser number of DoH servers may also contribute to a longer delay when compared to Non-DoH requests, which have a comparatively higher number of servers.

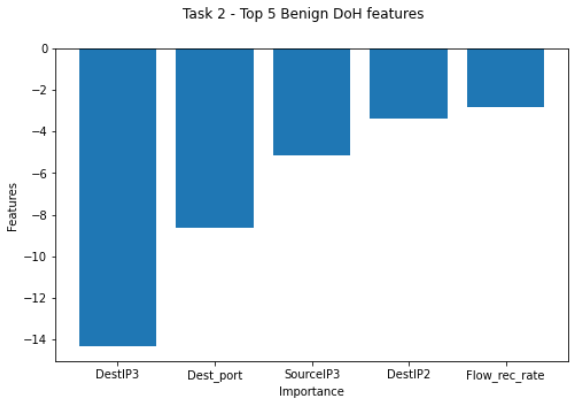

Fig 4c: Top 5 distinguishing features for Benign DoH

Figure 4c consists of the top 5 features that determine a DoH request to be a benign one. As DoH operates only on port 443 (same as HTTPS), and only a limited number of servers support DoH (like google, Cloudflare, Quad9), hence the dependency on the destination port is justified.

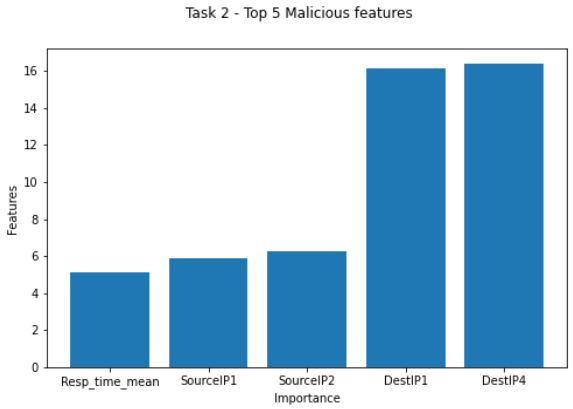

Fig 4d: Top 5 distinguishing features for Malicious DoH

Figure 4d consists of the top 5 features that determine a DoH request to be a malicious request. The DestinationIP4 of Malicious requests is predominantly '11', whereas the very few requests of benign DoH requests are made from '11'.

6. Conclusion

We use a simple machine learning model to perform two classification tasks. The first one is to classify the network traffic as DoH vs Non-DoH, and the second one being the classification of Benign DoH vs Malicious DoF. Fine inferences were drawn from the results and the highly impacting features were explained and justified for each class.

7. Reference

[1] Yaser M. Banadaki, "Detecting Malicious DNS over HTTPS Traffic in Domain Name System using Machine Learning Classifiers." Journal of Computer Sciences and Applications, vol. 8, no. 2 (2020): 46-55. doi: 10.12691/jcsa-8-2-2.

[2] Montazeri Shatoori, M., Davidson, L., Kaur, G., Habibi Lashkari, A., "Detection of DoH Tunnels using Time-series Classification of Encrypted Traffic," in The 5th IEEE Cyber Science and Technology Congress, Calgary, Canada, 2020.