Wannabe Programmers

Hackathon Serres 2019

Data Mining

SQL Injection ()
XSS ()
Local File Inclusion ()

Detect certain chars in the request.

Parse Data

Using the Data Mining we labeled each request as:

OKSQLIXSSLFI

SQL DataBase

Id	remote_host	request_url	request_type
1	220.243.135.5	/api/v1/login/?username=admin&password=1234	OK
2	62.109.16.162	/loain.php/?id=0%20or%201=1	SOLI

id	ip	country	totalRequests	code
1	54.36.149.94	France	2724	FRA
 2	45.61.164.120	United States	4055	USA

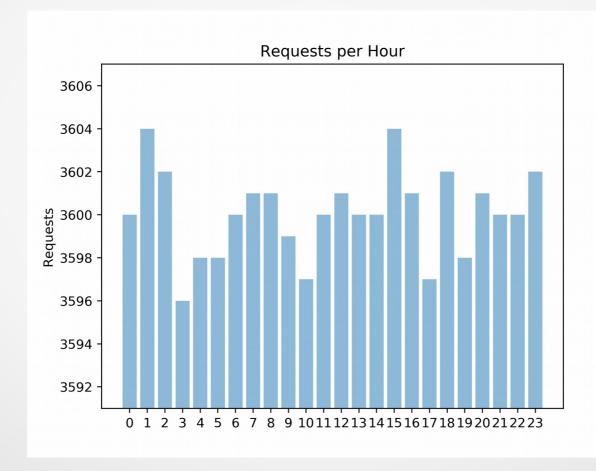
Answer to Questions

Most dangerous ip in our opinion:

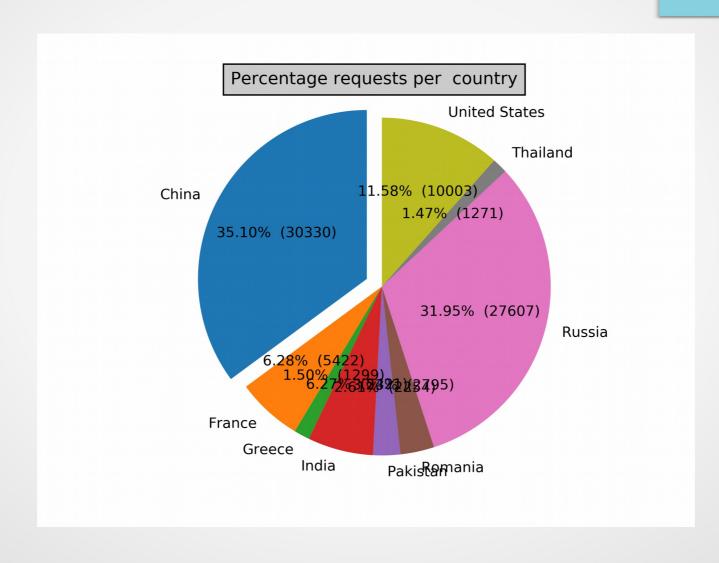
We used a function to evaluate the dangerousness of an ip, based on the number of each type of request.

Visualization

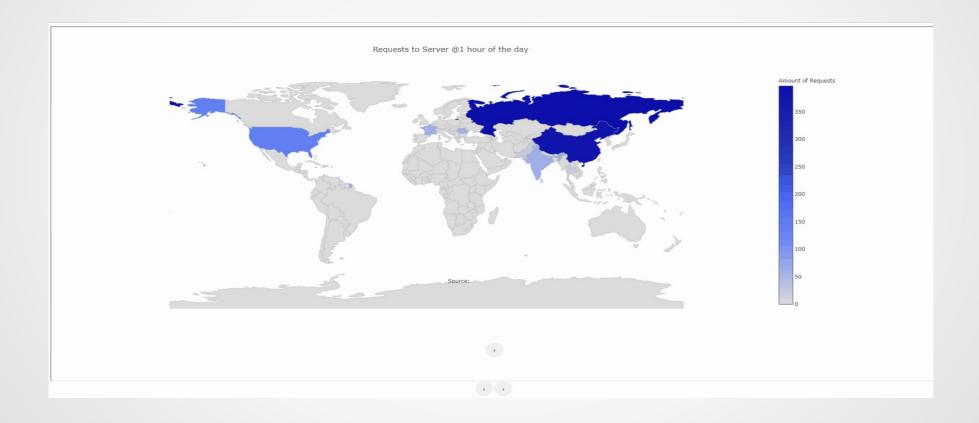
Requests per hour



Visualization



Visualization Word

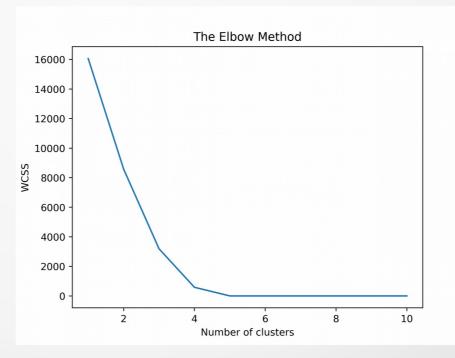


Site: https://bit.ly/2Js7Q4H

Machine Learning

Kmeans Clustering

Strings → 1,0 arrays based on some specific characters Sum of Squares



Machine Learning

Random Forest Classifier

Request → array of integers (unicode of chars)
Used 5-Fold cross validation:

Accuracy 99%