DataMining Project

Joshua García Sergi Carmona Óscar Conejo

01/05/2016

En el siguiente documento se realizará un análisis de las funcionalidades y capacidades de deteccion del antivirus perimetral de la empresa **JSO S.A.**

Se analizarán al detalle los siguientes puntos:

- **Ataques recibidos por franja horaria**
- **Ataques recibidos por pais de origen**
- **Ataques recibidos por origen de aplicación**
- **Ataques recibidos por nivel de criticidad**
- **Efectividad del sistema de reputacion web del antivirus**

Material:

- **Link download CSV file MEGA link (https://mega.nz/#!3t4VwYyb!YPSYUmgndXMGByXlGRsQJtkBp_-KkWYZPiUUX3wJTGU)**

Fuentes:

- **To learn more, see Github link (https://github.com/TheMaphius/DataMiningProject/tree/joshuagp).**

```
  setwd('..')
  if(!file.exists(paste( getwd(), "/resources", sep = ""))){
     dir.create("./resources")
  }

  path = paste(getwd(),"/resources/full_log_v2.csv",sep="")
  csv<<-read.csv(file=path, header=FALSE, sep=";", col.names = c("c1","TimeStamp","c
3","c4","c5","c6","c7","IP_origen","IP_Publica","c10","IP_interna","c12","c13","Protoco
lo","c15","Desde","Hacia","c18","c19","c20","c21","c22","c23","c24","Puerto_destino","c
26","c27","c28","c29","c30","c31","32","Malicioso","Nivel_de_riesgo","35","c36","c3
7","Pais_origen","Pais_destino","c40","c41","c42","c43","c44","c45","c46","c47","c4
8","c49","Sender","Subject","Remitente","c53","c54","c55","c56","c57","c58","c59","c6
0","c61","c62"), colClasses = c("character", "factor", "factor", "character" )) [,c("Ti
meStamp","IP_origen","IP_Publica","Protocolo","Desde","Hacia","Puerto_destino","Malicio
so","Nivel_de_riesgo","Pais_origen","Pais_destino","Sender","Subject","Remitente")]

  numLines<<-length(csv$TimeStamp)
```

Pregunta 1: Ataques recibidos por franja horaria

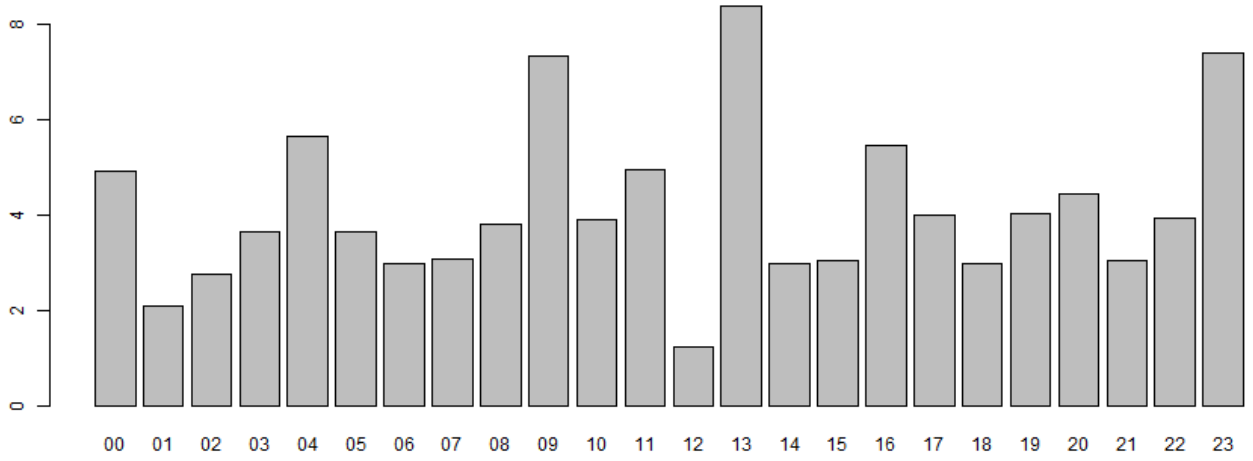
A continuación se muestra un gráfico mensual de los ataques recibidos por franja horaria.

El estado actual del gráfico, es una relacion de todos los ataques recibidos sin filtrar por mes y por año.

A modo de **ejemplo**, si filtramos por el mes de Marzo en el año 2016 se puede apreciar que a las 09:00 h se detectaron el mayor numero de ataques.

Month: All

Year: All



Pregunta 2: Ataques recibidos por pais de origen

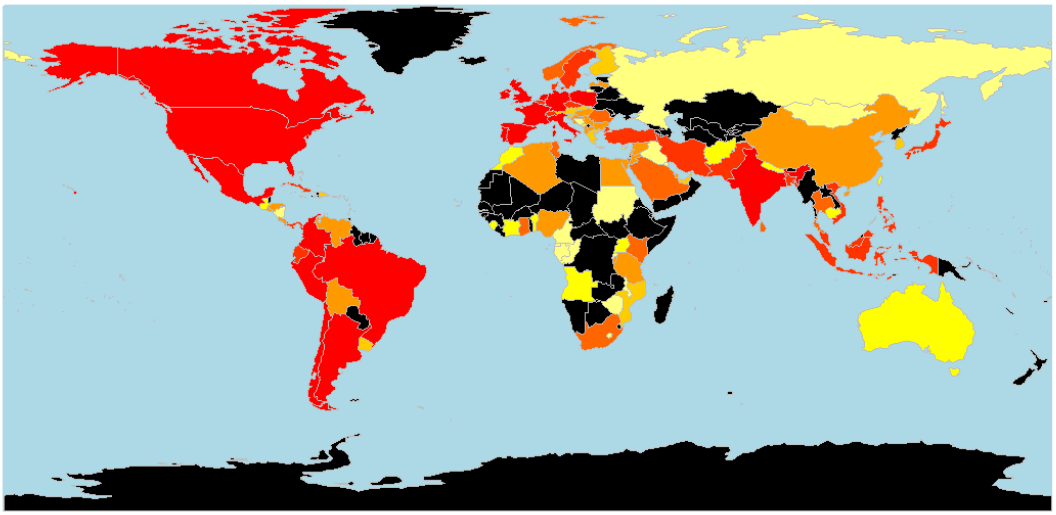
En el siguiente mapa se muestra el total de ataques recibidos por pais de origen.

A modo de leyenda, se representan los ataques con un rango de colores que varían desde:

Rojo: País originario del mayor número de ataques.

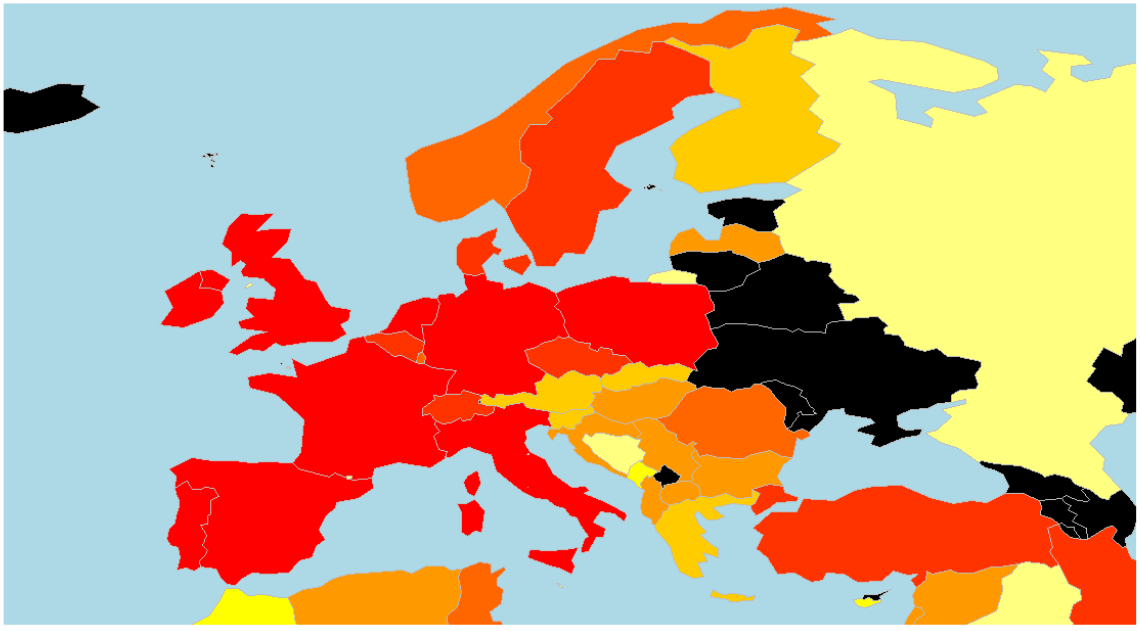
Amarillo: País originario del menor número de ataques.

World



El gráfico mostrado muestra los ataques a nivel mundial hacía nuestra empresa. Como se puede observar los continentes que más ataques originan son América y Europa.

Europe

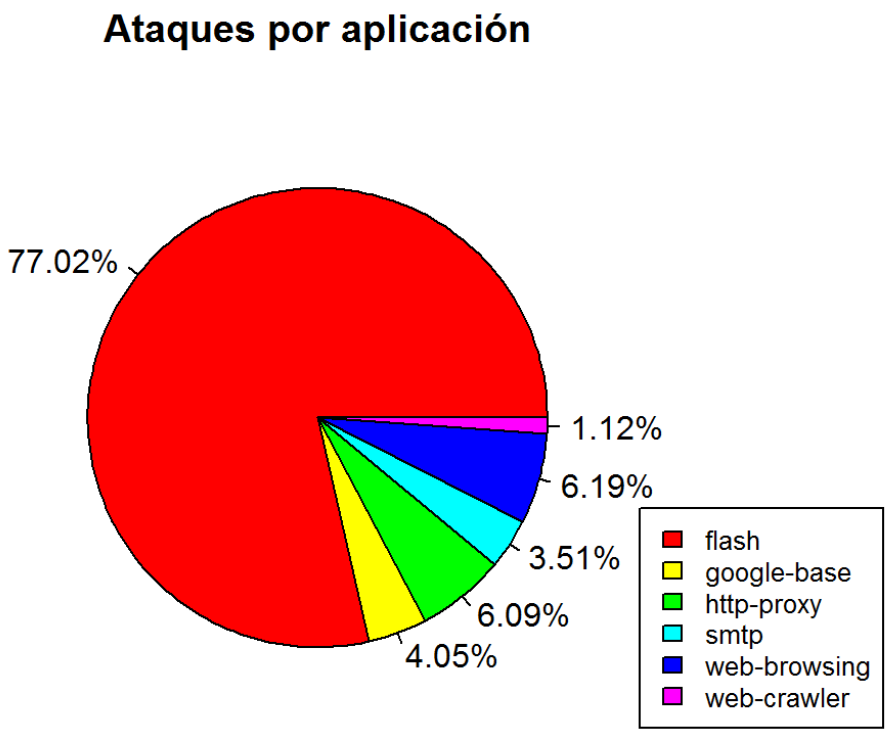


A nivel Europeo se puede ver como los países como España, Francia, Reino Unido, Alemania, Italia y parte de los países Nordicos inician ataques hacia nosotros que el antivirus puede parar.

Pregunta 3: Ataques recibidos por origen de aplicación

En el siguiente gráfico se muestran los ataques recibidos por origen de aplicación, es decir que protocolo/aplicación han dedicido utilizar los atacantes para vulnerar nuestro sistema.

Por ejemplo, **smtp**, **ftp**, **flash**, **web_browsing**, etc



Pregunta 4: Ataques recibidos por nivel de criticidad

En el siguiente gráfico muestra el número de ataques recibidos por nivel de criticidad, representado en porcentajes.

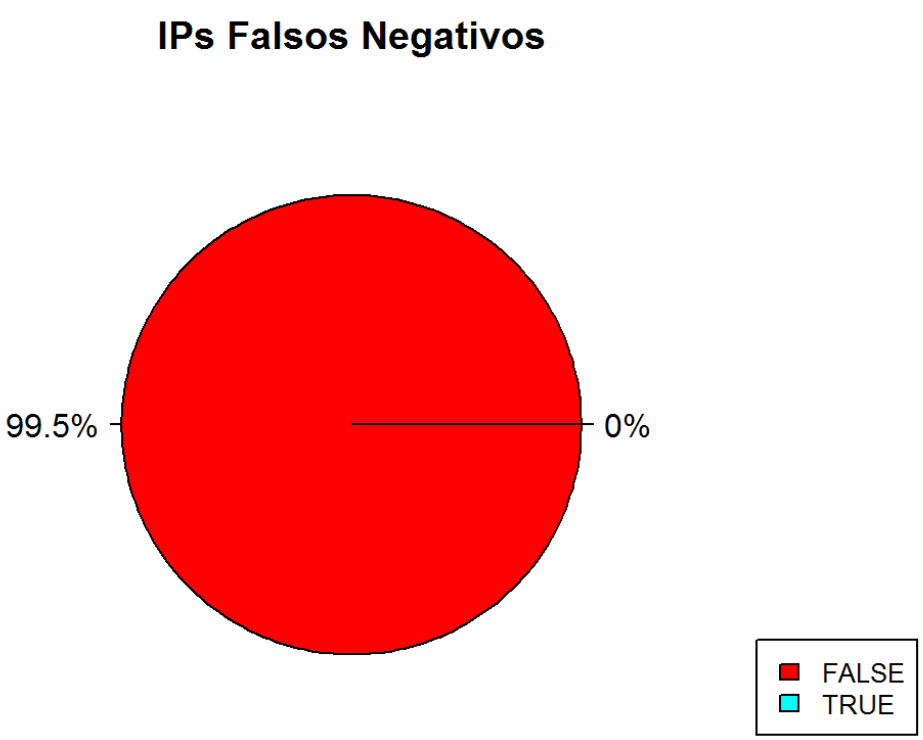
Los niveles de criticidad varian desde **critical** hasta **low**, donde **critical** es el nivel de ataque más dañino si logran vulnerar el sistema y **low** es el nivel de ataque con menos afectación en el sistema.

Pregunta 5: Efectividad del sistema de reputacion web del antivirus

En el siguiente gráfico se muestra el nivel de efectividad del sistema de reputación web (**WRS**) del antivirus perimetral, es decir:

Se comparan las IP's categorizadas por el WRS como **benignas** con una lista pública de IP's maliciosas.

En el caso de aparecer alguna de las IP en la lista negra publicada, se considerará como un fallo de categorización del antivirus y se añadirán en un fichero de texto que será utilizado a posteriori por el firewall perimetral con el fin de bloquear las IP's que aparezcan en el fichero y así poder aumentar el nivel de protección de nuestra infraestructura.




```
## [1] "Las siguientes IPs son falsos negativos:"
```

```
## [1] "68.169.54.179"
## [1] "217.216.65.224"
```

El resultado que podemos extraer es que el antivirus tiene un nivel bajo de falsos negativos ya que el número de IPs categorizadas como benignas son tal como dice ser.