23/04/2021

Miembros del grupo

Iago Pallares Tato (iago.pallares@udc.es)
Marcos Vázquez Campos (marcos.vazquez3@udc.es)
Daniel Osama González Anwar (daniel.osama.gonzalez@udc.es)

Resumen

La aplicación pretende ser una herramienta de utilidad para administradores de sistemas, busca ofrecer de forma simple un servicio que permita obtener información de la red y del tráfico que circula por la misma, a través del análisis de paquetería (ficheros pcap).

Listado exhaustivo de las funcionalidades a implementar

- Extracción de las IPs empleadas en la conexión.
- Filtrado de los paquetes por IP y protocolo.
- Cálculo de estadísticas: protocolos más usados, IPs más consultadas, fuentes de transmisión, protocolos usados, URLs más consultadas, etc.
- Generación de gráficos para visualizar las estadísticas, gráficos de sectores para mostrar el total de protocolos y IPs, histogramas con el tamaño de los payloads.
- Extracción de los dominios consultados por DNS.
- Geolocalización de IPs.
- Visualización en mapa de las IPs geolocalizadas
- Filtrado de tráfico, solo de las IPs locales.
- Detección de user agents maliciosos en trafico HTTP y análisis del tamaño de los payloads para buscar comportamientos anómalos.

Bocetos de las pantallas de la aplicación

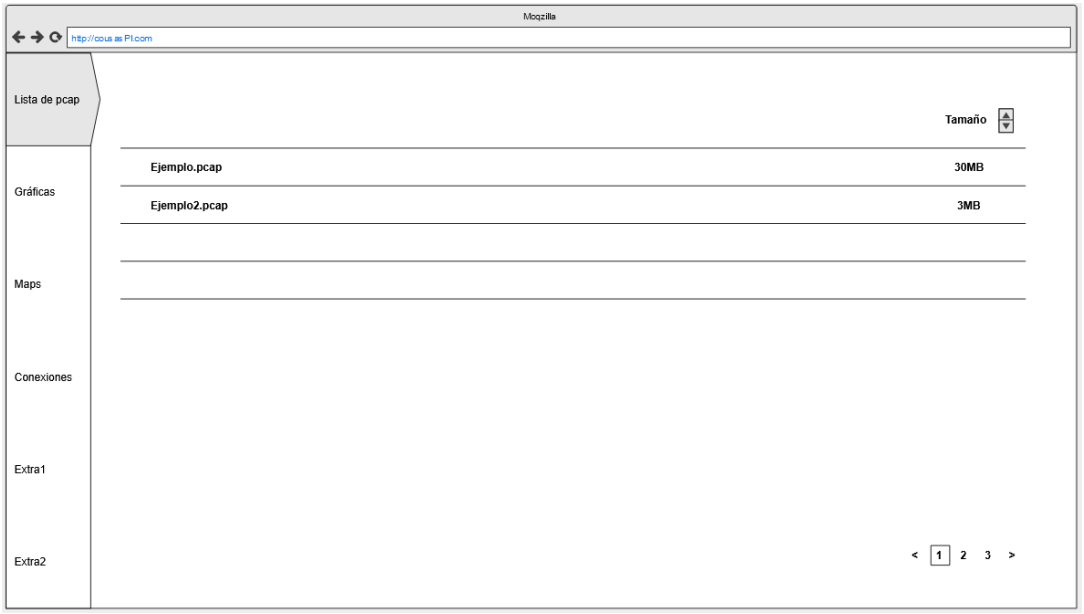

Figura 1: Pantalla de selección de pcaps.

Figura 2: Pantalla de muestra de gráficos y estadísticas.

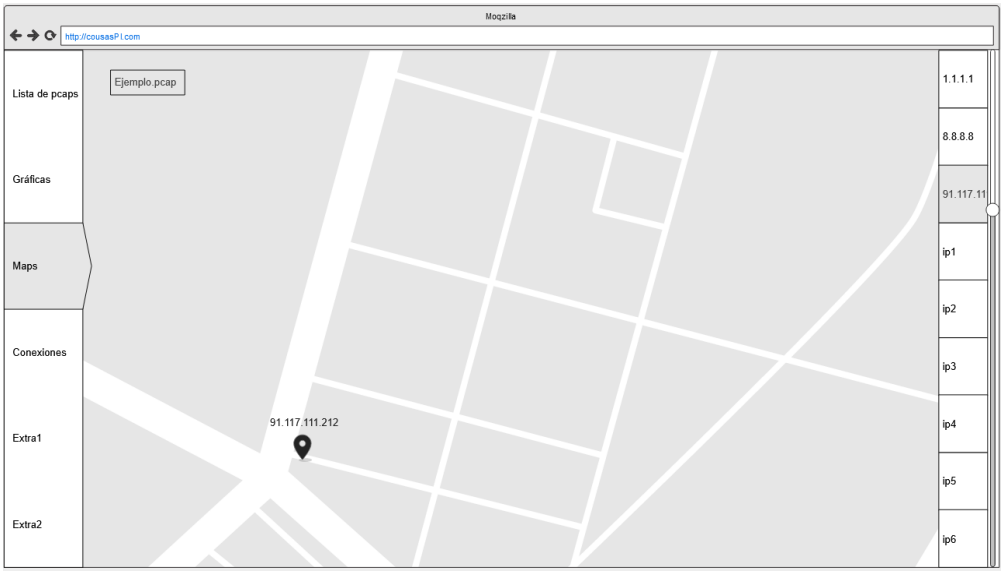

Figura 3: Pantalla de muestra de localización de una IP en el mapa.

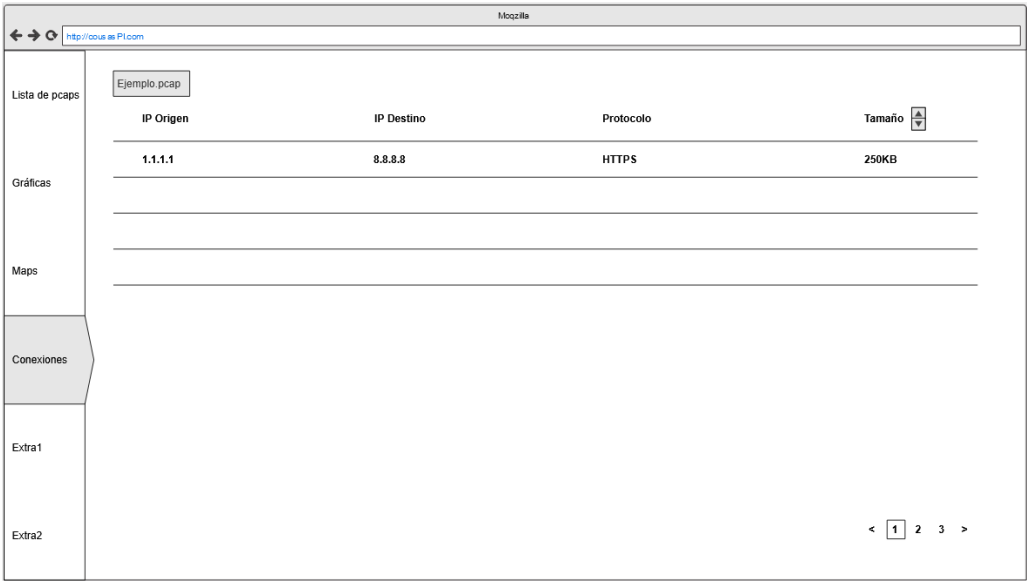

Figura 4: Pantalla de muestra de datos de red.

Flujo de datos de aplicación

Partimos de un fichero pcap, lo procesamos con Scapy y Pandas y obtenemos el dataframe, realizamos consultas, generamos tablas y gráficos. La información resultante puede ser tratada por las APIs que hemos seleccionado (para geolocalizar IP, mostrar en el mapa, dar información de la reputación de una URL, etc.).

APIs utilizadas

- API de Shodan para contrastar todo tipo de información de los datos previamente obtenidos (puertos, IPs, protocolos, etc.)
- API de Safe Browsing para consultar la reputación de URLs.
- API de apilayer ipstack para geolocalizar IPs.
- API de Google Maps, mostrar en el mapa a partir de la geolocalización de IPs.

Uso de Pandas dentro de la práctica

Conversión del fichero pcap a un dataframe, cuyos campos se corresponden, entre otros, con la ip de origen, ip de destino, puerto de origen, puerto de destino, protocolo, tamaño del payload.

Filtrado del dataframe originado mediante IP o protocolo.

Estadísticas obtenidas a partir del dataframe: IPs más consultadas, puertos más usados, etc.

Listar protocolos agrupados.

Detección de payloads anómalos mediante agrupación por protocolos y calculando la media del tamaño del payload para dicho protocolo, si alguno supera ampliamente dicha media conviene analizarlo.

Listado de todos los endpoints mostrando el par IP y MAC.

Uso de librerías adicionales de Python y otros lenguajes de programación

Uso de la librería de Python Scapy para realizar la inspección de elementos a nivel de bytes de los paquetes.

Uso de la librería Matplotlib para generar histogramas y gráficos a partir de los datos del dataframe que se obtiene del fichero pcap.

Generación de un grafo a partir de las IPs privadas del fichero pcap mediante JavaScript.

Consideraciones adicionales

Respecto a la generación del dataframe a partir del fichero pcap todavía no sabemos cómo haremos el procesado, habría varias opciones, realizarlo dinámicamente según vamos iterando los paquetes con Scapy o leer el dataframe a partir de un csv mediante el uso de tshark. Tendremos que analizar en su debido momento ventajas y desventajas cuando el proyecto esté más avanzado.

En cuanto a las funcionalidades a entregar para la primera iteración, debido a la potencial complejidad del proyecto, por ahora no tenemos muy claro qué funcionalidades vamos a tener disponibles para la primera entrega, pero en cualquier caso se espera contar con el proyecto de Django creado, consultas a la API de Shodan implementadas, la geolocalización y el núcleo del procesamiento a partir de Pandas hecho.

Es posible que en el futuro se añada una funcionalidad para hacer Login y que cada usuario tenga una lista personal de pcaps a analizar.