

Ana Paula Katsuda, A01025303
Andréina Sanáñez, A01024927
Programación de estructuras de datos
Prof. Jorge Rodríguez
19 de Noviembre 2021

Reporte Reto 4

Introducción

Analizar y procesar información sobre distintas interacciones de usuarios con distintos servicios en la computación, es inmensamente importante para identificar si existen accesos maliciosos en redes y proteger datos relevantes. Los accesos maliciosos tienen la posibilidad de afectar enormemente a empresas e individuos; puede costar la privacidad de éstos, e incluso puede tener impactos monetarios significativos.

Dada la importancia de dar seguimiento de dichas interacciones, el objetivo del presente escrito es continuar con el análisis de las interconexiones de los usuarios pertenecientes a la empresa “reto.com” con servicios externos a ella. En este caso, se expandirá lo realizado en otros escritos para determinar la cantidad de conexiones entrantes de los dominios destino y obtener los sitios más visitados cada día de registro.

Para hacer lo anterior, se utilizaron diccionarios con el fin de relacionar un dominio con su número de conexiones entrantes por día y se usó un árbol de búsqueda binario para determinar los más frecuentados.

Imágenes de los resultados obtenidos a través del código

```
Top 5 del número de conexiones en la fecha 10-8-2020:  
1369: protonmail.com,  
20: instagram.com,  
16: spotify.com,  
16: merriam-webster.com,  
16: dictionary.com,
```

```
Top 5 del número de conexiones en la fecha 11-8-2020:  
1336: protonmail.com,  
20: ca.gov,  
19: cbssports.com,  
18: wayfair.com,  
18: usps.com,
```

```
Top 5 del número de conexiones en la fecha 12-8-2020:  
1319: protonmail.com,  
21: wellsfargo.com,  
20: cnet.com,  
19: bankofamerica.com,  
18: theguardian.com,
```

```
Top 5 del número de conexiones en la fecha 13-8-2020:  
1331: protonmail.com,  
22: gamepedia.com,  
19: npr.org,  
19: mapquest.com,  
17: macys.com,
```

```
Top 5 del número de conexiones en la fecha 14-8-2020:  
1368: protonmail.com,  
21: bankofamerica.com,  
19: hulu.com,  
18: wiktionary.org,  
17: quora.com,
```

```
Top 5 del número de conexiones en la fecha 15-8-2020:  
0: zillow.com,  
0: youtube.com,  
0: yelp.com,  
0: yellowpages.com,  
0: yahoo.com,
```

```
Top 5 del número de conexiones en la fecha 16-8-2020:  
0: zillow.com,  
0: youtube.com,  
0: yelp.com,  
0: yellowpages.com,  
0: yahoo.com,
```

```
Top 5 del número de conexiones en la fecha 17-8-2020:  
1162: protonmail.com,  
303: p7necksgkbynlvmgenv5.org,  
15: washingtonpost.com,  
15: msn.com,  
14: steamcommunity.com,
```

```
Top 5 del número de conexiones en la fecha 18-8-2020:  
1018: protonmail.com,  
611: p7necksgkbynlvmgenv5.org,  
208: wiktionary.org,  
15: apartments.com,  
14: pinterest.com,
```

```
Top 5 del número de conexiones en la fecha 19-8-2020:  
1100: protonmail.com,  
599: p7necksgkbynlvmgenv5.org,  
18: cnn.com,  
17: macys.com,  
14: youtube.com,
```

Top 5 del número de conexiones en la fecha 20-8-2020:

1071: protonmail.com,
593: p7necksgkbynlvmgenv5.org,
17: aol.com,
15: foxnews.com,
15: en.wikipedia.org,

Top 5 del número de conexiones en la fecha 21-8-2020:

1123: protonmail.com,
588: p7necksgkbynlvmgenv5.org,
16: irs.gov,
15: wiktioary.org,
15: investopedia.com,

1. *¿Existe algún sitio que se mantenga en el top 5 todos los días?*

Es posible visualizar a partir de las imágenes de lo obtenido mediante el código, que el sitio que se mantiene en el top 5 (e incluso siendo el primer lugar), es “protonmail.com”. Esto implica que los usuarios en general, utilizan constantemente este sitio durante los días de trabajo. Resulta interesante notar que durante los días 15-8-2020 y 16-8-2020 no existieron conexiones; lo anterior puede deberse a que estos días eran fin de semana.

2. *¿Existe algún sitio que entre al top 5 a partir de un día y de ahí aparezca en todos los días subsecuentes?*

Efectivamente, el dominio “p7necksgkbynlvmgenv5.org” aparece a partir del día 17-8-2020 hasta el 21-8-2020, que es la última fecha que se registra en dicha base de datos. Asimismo, cabe mencionar que este es uno de los dominios anómalos que se identificó en la práctica anterior, y que como se puede observar tiene un número considerable de conexiones entrantes por día.

3. *¿Existe algún sitio que aparezca en el top 5 con una cantidad más alta de tráfico que lo normal?*

Es notorio que tanto “protonmail.com” como “p7necksgkbynlvmgenv5.org” tienen cantidades de tráfico extremadamente altas. Por un lado, se puede observar que la conexiones entrantes del dominio “protonmail.com” se mantiene relativamente constante (alrededor de 1300 conexiones por día) y se encuentra presente durante todos los días registrados. Ahora bien, en este caso dicho comportamiento es normal principalmente dado que “protonmail.com” es un dominio que pertenece a un servicio conocido, y además es de esperarse que los usuarios de la empresa reto.com tengan múltiples conexiones al servicio de email durante cada día.

Por otro lado, “p7necksgkbynlvmgenv5.org” no solo es un dominio anómalo, sino que la presencia de este no tomaba lugar en la actividad de conexiones de la compañía durante los primeros días registrados en la base de datos. Específicamente, el número de conexiones del mismo se encuentra en un rango de las 300 - 600 aproximadamente.

Aportaciones Individuales

En cuanto a las aportaciones, es importante mencionar que el análisis de la situación problema, así como su resolución fue hecha en conjunto por las dos integrantes del equipo. Si bien ciertas partes fueron realizadas de manera individual, se tuvieron varias sesiones de revisión para verificar el correcto funcionamiento de la incorporación de cada una de las

aportaciones. A continuación, se mencionan las divisiones de tareas realizadas por cada integrante:

En conjunto:

- Análisis del problema.
- Pruebas del funcionamiento del código y de cada uno de los incisos.
- Reporte.

Andreína Sanáñez / A01024927

- Implementación de la función conexionesPorDia().

Ana Paula Katsuda / A01025303

- Implementación de función top().