Protocolo ARP

September 23, 2014

Universidad de Buenos Aires - Departamento de Computación - FCEN

Integrantes:

- Gallardo, Guillermo L.U.: 32/10 gagdiez.c@gmail.com
- Fixman, Martin L.U.: 391/11 martinfixman@gmail.com
- Matayoshi, Leandro L.U.: 79/11 leandro.matayoshi@gmail.com
- Szyrej, Alexander L.U.: 642/11 alexander.szyrej@gmail.com

Contents

1 Introducción

El protocolo ARP (*Adress Resolution Protocol*) permite mapear direcciones IP a direcciones físicas. Cuando un host desea comunicarse con otro dentro de una misma red local debe hacerlo utilizando la *MAC Adress* del mismo. Conociendo su dirección IP, puede *brodcastear* un paquete *ARP who-has* pidiendo la *MAC* asociada a dicha IP. De esta manera recibirá en respuesta un paquete *ARP-reply* por parte del host destino. Para evitar una cantidad excesiva de broadcasts, cada nodo de la red mantiene su propia tabla de mapeos que actualiza luego de cierto tiempo.

Cada paquete ARP incluye numerosos campos. Sin embargo, pondremos énfasis especialmente en los siguientes:

- Operation: 1 (Request), 2 (Reply)
- Sender hardware adress: *MAC* address source
- Sender protocol adress: *IP* address source
- Target hardware adress: *MAC* address dest
- Target protocol adress: *IP* address dest

El objetivo de este trabajo práctico es analizar el flujo de paquetes *who-has* en distintas redes no contraladas. Luego, en base a la información recopilada, analizaremos el rol que juega cada dispositivo dentro de la red. De esta manera podemos corroborar de manera empírica el comportamiento establecido por el protocolo *ARP* para el envío de paquetes en una red local.

2 Desarrollo

El primer paso en este trabajo fue implementar un script en *python* para escuchar pasivamente paquetes dentro de la red local. Para esto capturamos todos los paquetes que circulan utilizando la librería Scapy, y nos quedamos solo con los ARP de tipo *who-has*. Luego, contamos cuantas veces aparece una dirección en cada campo, así como también la interacción entre todos los pares de ip. Terminamos ordenando esta información según la cantidad de paquetes, de mayor a menor.

Con estas observaciones nos proponemos a calcular la entropia tomando dos modelos distintos de fuente: en S_{src} los símbolos son las direcciones IP que aparecen en el campo SRC; en S_{dst} son las direcciones IP que aparecen en el campo DST.

En relación a esto tenemos una mejor base para realizar un análisis del estado de la red al momento de las mediciones, y podemos analizar estadísticamente qué IPs son más relevantes o significativas en la LAN utilizando la información del símbolo con respecto a la entropía de la fuente correspondiente.

En la próxima sección podrán verse los datos obtenidos y el análisis de los mismos para las distintas redes en las cuales fueron realizadas las capturas: Entrepiso DC (**2634** paquetes), Fibertel Zone de Starbucks (**500** paquetes), Casa de Familia (**515** paquetes) y Empresa (**400** paquetes).

3 Resultados y análisis

3.1 Casa Familia

En el caso de la casa familiar sabemos que la dirección del router es 192.168.1.1 y que 192.168.1.35 pertenece al dispositivo que más tiempo estuvo prendido, el mismo es una notebook. Podemos apreciar que la ip del router es de las que aparece con mayor frecuencia en los campos SRC (fig *a*) o DST (fig *b*), mientras que la notebook solo aparece un alto número de veces en DST.

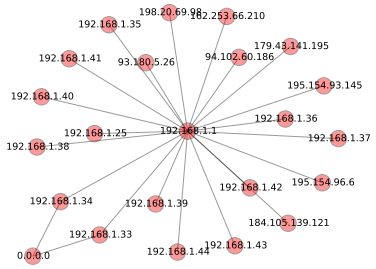
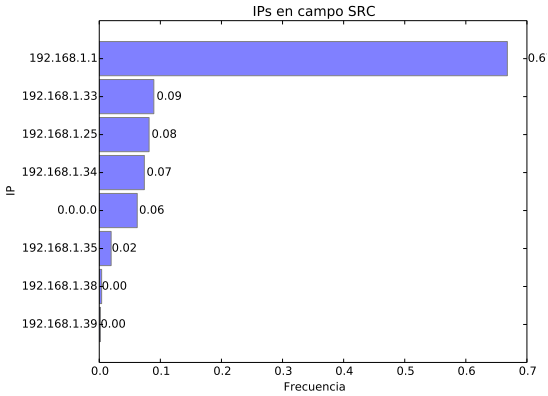


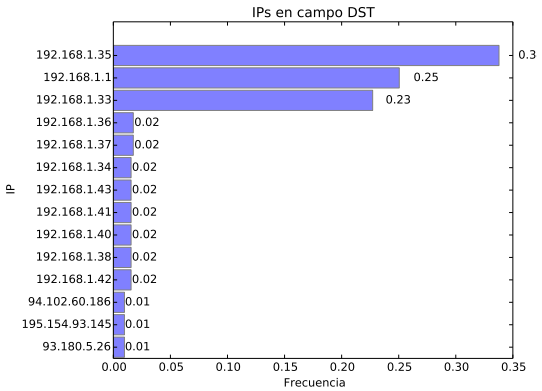
Figure 1: Grafo de la red Casa

En la *Figura 1* podemos ver que la red local se puede representar como un grafo en forma de estrella, el mismo tiene al router como centro.

La entropía fue: 1.68 para el modelo *SRC* y 2.67 para el modelo DST. Cabe destacar que aunque no está reflejado en los gráficos, sucedió algo interesante, aparecieron paquetes ARP con direcciones que no pertenecian a la red local.



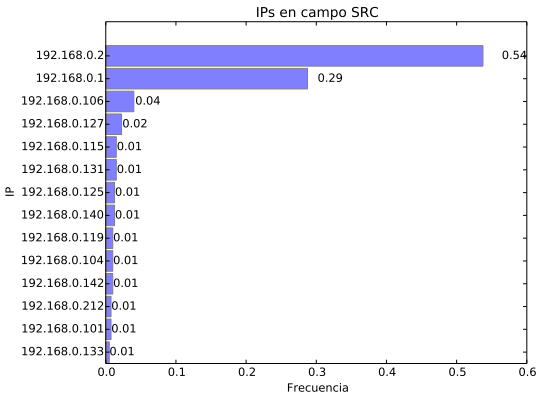
Estimación de la probabilidad de cada símbolo en modelo SRC



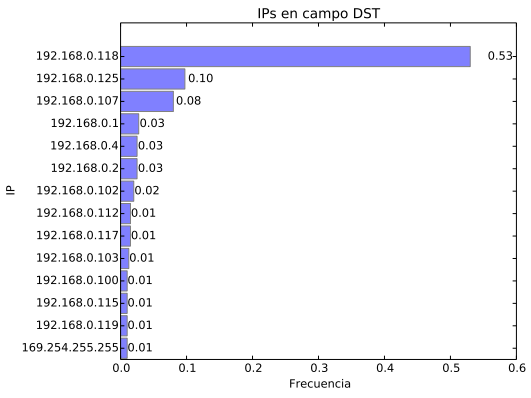
Estimación de la probabilidad de cada símbolo en modelo DST

3.2 Empresa

Nuevamete, en esta red teníamos conocimiento de quien era el router y las computadoras mas importantes. Lo primero que notamos en este caso fue que la topología del grafo de la red (*Figuras 4 y 5*) no coincide con la de una estrella. Es posible observar una interacción mucho mayor entre distintos pares hosts, y no únicamente entre cada uno de estos y el router.



Estimación de la probabilidad de cada símbolo en modelo SRC



Estimación de la probabilidad de cada símbolo en modelo DST

En el gráfico de barras del modelo SRC el router juega un papel relevante ya que es el segundo símbolo con mayor cantidad de emisión de paquetes. Es interesante mencionar que el nodo con mayor cantidad de paquetes enviados, cuya ip es 192.168.0.2, corresponde a un servidor apache. El resto de los nodos que aparecen en el gráfico corresponden a notebooks y celulares conectados a la red.

En el gráfico de barras del modelo DST es en donde se nota la mayor cantidad de diferencias. El router no se corresponde con ninguno de los 3 nodos con mayor cantidad de paquetes recibidos:

- 192.168.0.118 corresponde a una notebook
- 192.168.0.125 corresponde a un servidor apache
- 192.168.0.107 corresponde a un servidor apache

Recién aparece en la cuarta posición, con una cantidad de paquetes recibidos considerablemente menor a los primeros. En esta red en particular los dispositivos se envían paquetes who-has entre ellos en mayor medida, y hacia los servidores en vez de hacia el router.

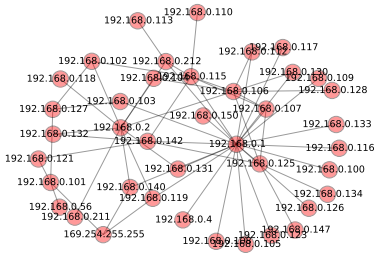


Figure 4: Topología de la red en una empresa

El valor de la entropía para el modelo SRC es 2.05. El valor de la entropía para el modelo DST es 2.99

Los gráficos de barra muestran que el router ha perdido preponderancia respecto a la emisión y recepción de paquetes en comparación con otros nodos distinguidos.

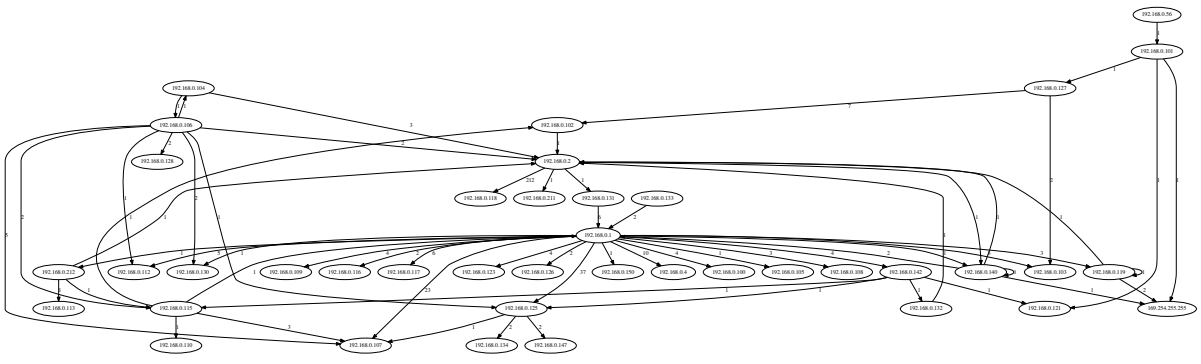


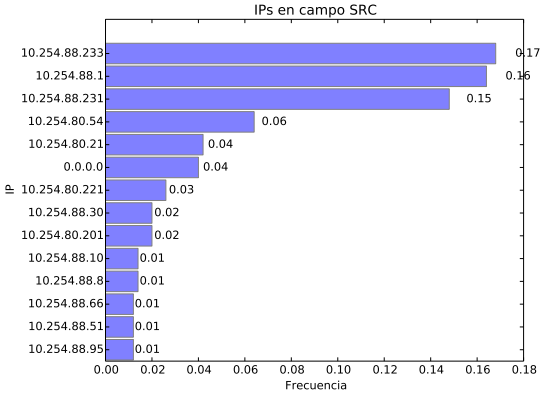
Figure 5: Grafo de la red Empresa

3.3 Starbucks

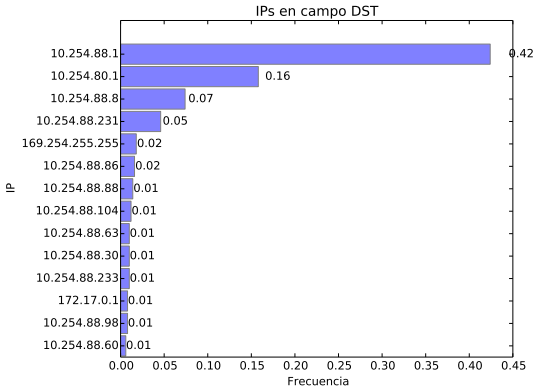
Para el caso de la red abierta disponible en Starbucks, podemos ver en la *Figura 7* una notoria centralización de paquetes hacia/desde la dirección 10.254.88.1, la cual a su vez posee una alta frecuencia de aparición en ambos modelos. Esto sería el comportamiento esperado del router de la red 10.254.88.0/24. Asumimos que ésta es la dirección de la red puesto que la mayoría de las IPs de los paquetes capturados difieren en el último octeto. Por otro lado, podemos ver que la segunda dirección con mayor frecuencia en DST fue 10.254.80.1, seguida por 10.254.88.8. Suponemos que 88.8 es la pc del lugar dada la ip baja, y la cantidad de apariciones. Por otro lado, no sabemos que es 80.1 dado que parece formar una red aparte de la sniffeada.

La entropía fue: 4.61 para el modelo *SRC* y 3.76 para el modelo DST.

Nuevamente entre los paquetes capturados volvieron a aparecer direcciones que no parecieran pertenecer a la red local, resultando sumamente interesante el caso de 10.254.80.1



Estimación de la probabilidad de cada símbolo en modelo SRC



Estimación de la probabilidad de cada símbolo en modelo DST

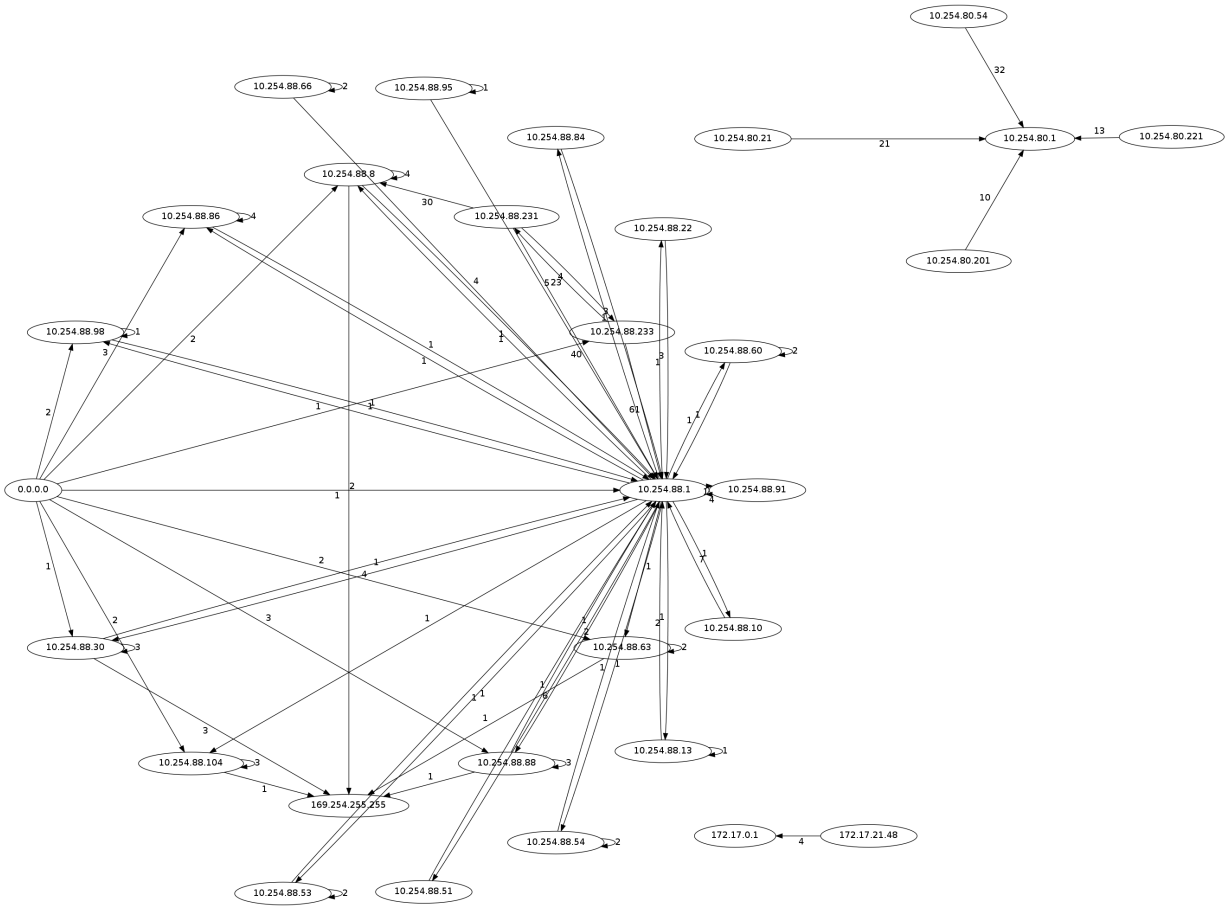


Figure 7: Grafo de la red Starbucks

3.4 Entrepiso

En el caso del entrepiso las *Figuras 8 y 10* muestran tres nodos con la forma que caracterizó al router dentro de la casa, estos nodos son: 10.1.100.254, 10.1.200.30 y 10.1.200.254. Sin embargo, estos poseen al menos la mitad de las aparaciones que otros nodos en el modelo DST, cosa que 200.30 repite en SRC. La entropía fue: 4.25 para el modelo *SRC* y 4.90 para el modelo DST. El Entrepiso fue la red con mas entropía. Una vez mas aparecieron direcciones sueltas, como es por ej: 10.1.11.254.

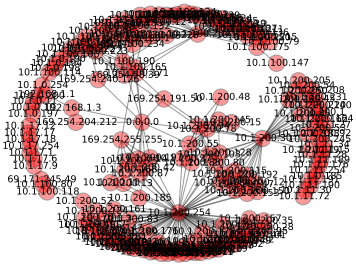
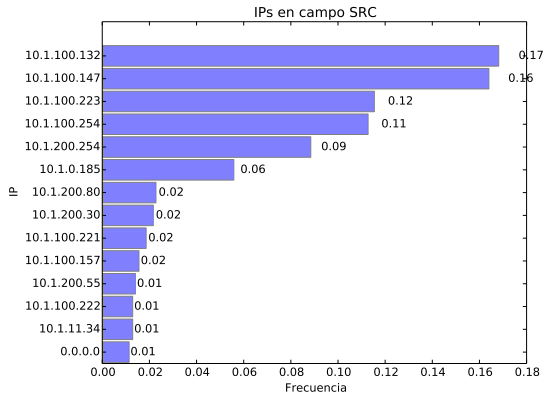
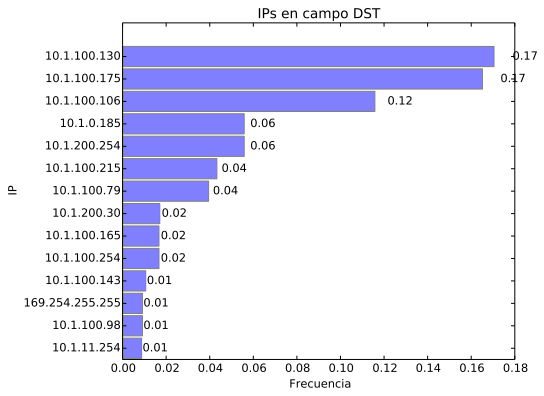


Figure 8: Grafo de la red Entrepiso



Estimación de la probabilidad de cada símbolo en modelo SRC



Estimación de la probabilidad de cada símbolo en modelo DST

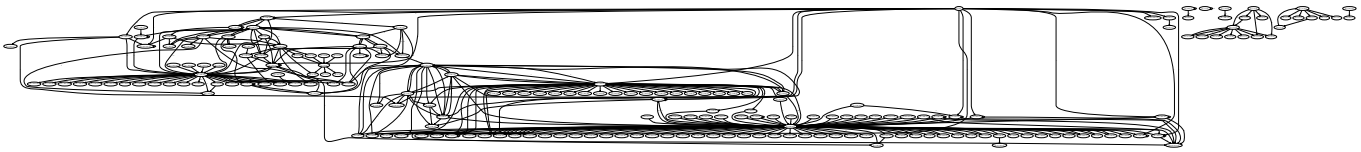


Figure 10: Digrafo de conectividad en el entrepiso

3.5 Discusión

En todas las redes sucedió que aparecieran direcciones ip que
se supone no pertenecen a la misma.

En algunos casos apareció la dirección 169.254.255.255. Investigando encontramos que es utilizada como broadcast por DHCP que es un protocolo de configuración automática de parámetros de red tales como direcciones IP para interfaces y servicios.

La dirección 0.0.0.0 es el estándar para *broadcastear* dentro de una red local.

Por otra parte, el resto de las direcciones IP que aparecen pueden deberse a que los routers posean la opción de *Proxy ARP* habilitada, cosa que tiene mucho sentido en el caso del Entrepiso por ejemplo.

A su vez, si bien en los grafos se notó claramente la dirección ip de los routers, ésto sólo se vio reflejado en la estimación de la probabilidad en los modelos únicamente en la Casa y el Starbucks. Ésto muestra que el contar la cantidad de apariciones de una IP en los campos no es suficiente para detectar los routers de la red.

Es posible que las direcciones que aparecen con mayor frecuencia en el grafo de pedidos ARP correspondan los dispositivos que estabam más activos durante el periodo de tiempo que se hizo el sniffing.

4 Conclusiones

Vimos que capturando paquetes who-has de una red se puede hacer cierto análisis de la misma y determinar de alguna forma la estructura de la red, reconociendo IPs de mayor interacción a nivel ARP como fue el caso de routers o servidores. Así mismo se descubren IPs que, en principio, uno no esperaría que esten involucradas en absoluto, como se ha visto anteriormente.

Cuando se analiza una red pequeña, por ejemplo, una LAN hogareña, el $sniffeo$ de estos paquetes permite un analisis quizá más detallado. Los dispositivos conectados a la red se comunican principálmente con el router y al no ser demasiadas las IPs involucradas, los grafos son pequeños y bastante claros, del tipo estrella centralizados en el router.

En el caso de otras redes más grandes, la conectividad empieza a enroscarse y análisis se complica. Ya los routers no son tan claros de encontrar, e IPs de otros dispositivos (ej: servidores) también toman relevancia,

quizás en mayor medida que un router.

5 Referencias

- **Computer Networks: A Systems Approach**, *Larry L. Peterson and Bruce S. Davie.*
- **Computer Networks**, *Andrew S. Tanenbaum*
- **Special-Purpose IP Address Registries** (RFC 6890), *Internet Engineering Task Force (IETF)*
- **An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware** (RFC 826), *David C. Plummer*