DEPARTAMENTO DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

TP1: Wiretapping

Teoría de las Comunicaciones

Integrante	LU	Correo electrónico
Furman, Damián	936/11	damian.a.furman@gmail.com
Lambrisca, Santiago	274/10	santiagolambrisca@hotmail.com
Marottoli, Daniela	42/10	dani.marottoli@gmail.com
Vanecek, Juan	169-10	juann.vanecek@hotmail.com

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - (Pabellón I/Planta Baja)
Intendente Güiraldes 2160 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (54 11) 4576-3359
http://www.fcen.uba.ar

Índice

1. Introducción

Aprovechando las herramientas existentes para el análisis de transferencia de paquetes, como Scapy y Wireshark, nos desarrollamos nuestra propia herramienta que nos permite captar los paquetes de la redes local a donde estemos conectados. Para poder realizar esto, tuvimos que valernos de una modalidad de uso brindada por la placa de red.

Así, utilizando la placa de red en modo Promiscuo o Monitor, nos dispusimos a captar los paquetes correspondientes al protocolo ARP (Address Resolution Protocol), con el objetivo de realizar un análisis sobre el intercambio de paquetes de este protocolo realizado en distintas redes, buscando identificar los nodos más significativos e intentando comprender su rol dentro de la red. Este tipo de paquetes es adecuado para este análisis ya que en redes de acceso múltiple es el encargado de traducir direcciones de red (IP) en direcciones de enlace (MAC). Los hosts los envían cuando quieren conocer la ubicación de cierta IP, y un router está constantemente actualizando su tabla de routeo, por lo que podríamos identificar a éstos de acuerdo al flujo de ARPs que corren por la red.

Valiéndonos de distintas herramientas de análisis y graficación hemos realizado este trabajo, obteniendo los resultados presentados a continuación.

2. Desarrollo

En el primer punto se pide implementar una herramienta para escuchar pasivamente una red local. Scapy nos provee una serie de métodos como `sniff` que ejecuta un callback cada vez que la placa de red recibe un paquete. Luego la clase `Sniffer` se encarga de parsearlo si es un paquete ARP, y guardarlo convenientemente.

El paquete está compuesto, entre otras cosas, por la dirección IP y MAC origen y destino, y el tipo de consulta: *who-has* o *is-at*.

Como método para identificar los routers en la red analizamos tres fuentes de información en 5 redes diferentes (dos domésticas, una empresarial, una de un organismo, y una pública). Las fuentes que usamos fueron:

1. *IP origen*; evento: IP X origen de un paquete *who-has*.
2. *IP destino*; evento: IP X destino de un paquete *who-has*.
3. *IP origen - IP destino*; evento: IP X manda un paquete who-has a y.

Para cada una de estas fuentes, la clase `Sniffer` contiene un diccionario para almacenar cada evento.

Una vez que ya tenemos la estructura armada, pusimos a correr el programa unos 30 minutos en cada LAN, un tiempo que consideramos prudente para poder tomar conclusiones.

Como queremos encontrar los puntos distinguidos en la red, los vamos a considerar a partir de los eventos que posean menos información o, lo que es lo mismo, que tengan una mayor probabilidad de que suceda. En particular, nos interesan los eventos s que cumplan $I(s) - H(S) < 0$.

3. Gráficos y análisis

Para las distintas redes utilizadas, presentamos los datos obtenidos a través de distintos gráficos, uno para cada fuente de información considerada, que nos permiten, no solo mostrar mas claramente los resultados obtenidos, sino que también son útiles para realizar el análisis de las distintas redes y compararlas entre sí.

Consideramos de utilidad, para analizar las fuentes de información 1 y 2 presentadas en el desarrollo, representar la actividad mostrada por cada nodo dentro de una red graficando la cantidad de información proporcionada por el evento relacionado a este. Además decidimos mostrar una línea color rojo que representa la Entropía de la fuente. El echo de que la información brindada por un nodo se encuentre debajo de la linea roja, nos indica que ese nodo tiene una actividad significativa dentro de la red y nos dice que dicho nodo es de importancia a la hora de realizar el análisis.

Por otro lado, para el analisis de la fuente de información 3, presentamos un grafo, con nodos de distintos tamaños, donde los nodos representan una IP y los ejes un un paquete que lo referencia como destino o fuente según la dirección. El tamaño de los nodos representa la cantidad de intercambios en los que participó, viendose así, los nodos mas significativos graficados con mayor tamaño y siendo facilmente identificables.

3.1. Red WiFi casa particular 1

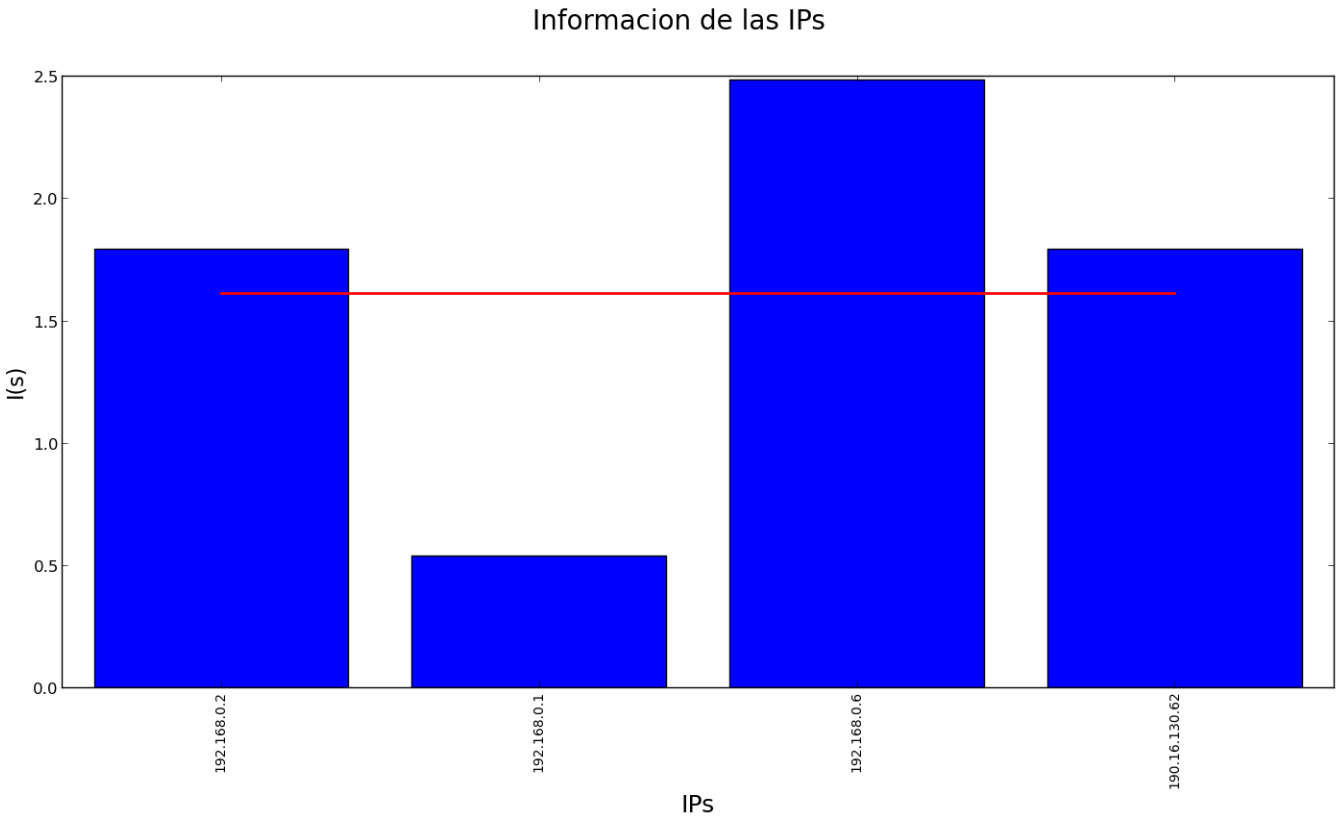

Figura 1: Fuente de información: IPs que envían

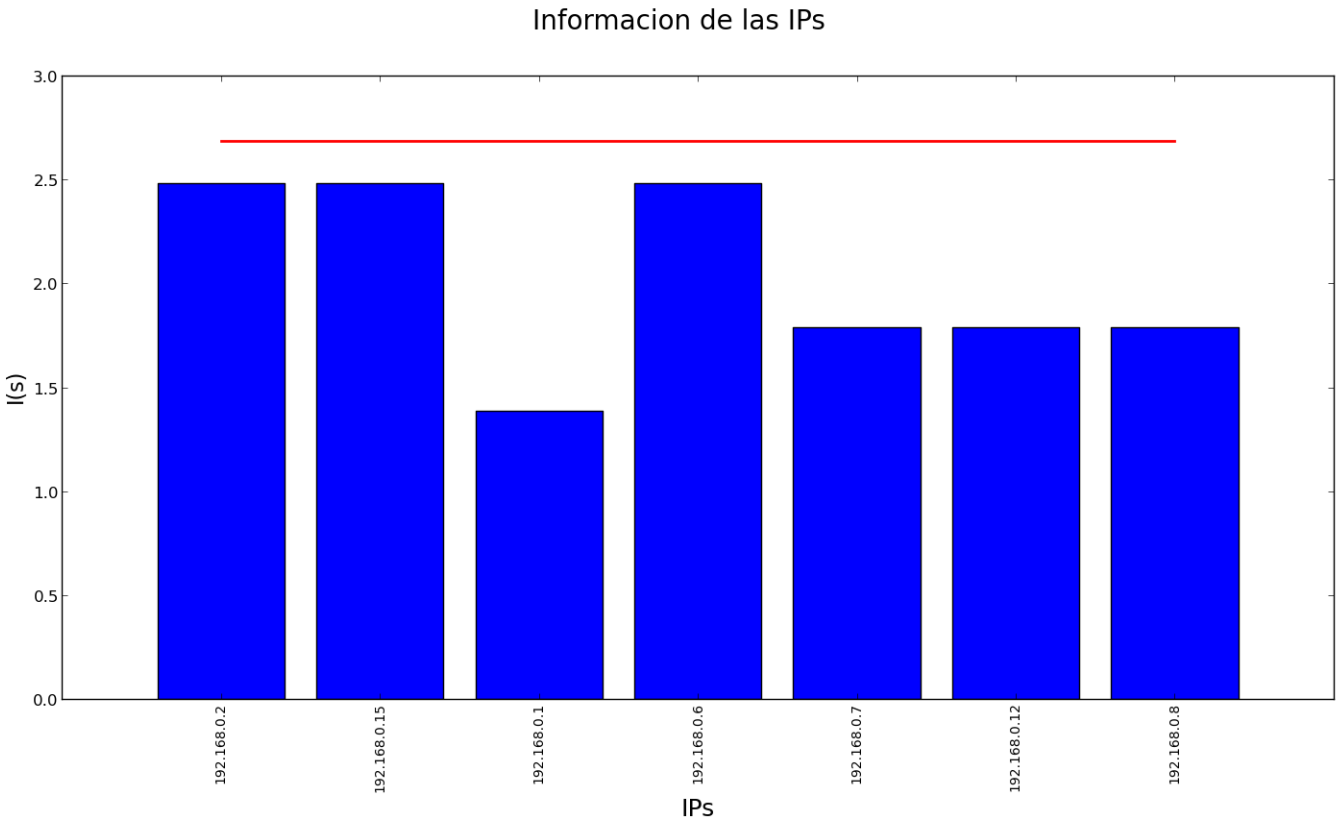

Figura 2: Fuente de información: IPs que reciben

¡¡¡¡¡¡¡ HEAD

Los gráficos presentados anteriormente pertenecen a una red doméstica. Podemos observar que uno de los nodos de la red representa una cantidad de unidades de información notablemente menor a la de los demás. Siendo que la función que determina las unidades de información para cada evento, es decreciente, mientras más veces haya ocurrido este evento en el periódo de tiempo que se tomó la muestra, menor será la cantidad de información que representa el hecho que ese evento ocurra. Así, vemos que el evento cuya información se destaca por ser notablemente menor a los demás, es el que más intercambio de paquetes ha realizado. Tentados a pensar que éste debería ser el nodo de la red que representa al Router, pudimos comprobarlo en la configuración de la red. Otro indicio fácil de notar había sido la dirección IP asociada a este nodo (192.168.0.1) ======= Los gráficos presentados anteriormente pertenecen a una casa particular. Podemos observar que uno de los nodos de la red representa una cantidad de unidades de información notablemente menor a la de los demás. Siendo que la función que determina las unidades de información para cada evento, es decreciente, mientras más veces haya ocurrido este evento en el periódo de tiempo que se tomó la muestra, menor será la cantidad de información que representa el hecho que ese evento ocurra. Así, vemos que el evento cuya información se destaca por ser notablemente menor a los demás, es el que más intercambio de paquetes ha realizado. Tentados a pensar que éste debería ser el nodo de la red que representa al Router, pudimos comprobarlo en la configuración de la red. Otro indicio fácil de notar había sido la dirección IP asociada a este nodo: 192.168.0.1 ¿¿¿¿¿¿¿ ca62944151b53781379a691d4d118fdec52d6fc4

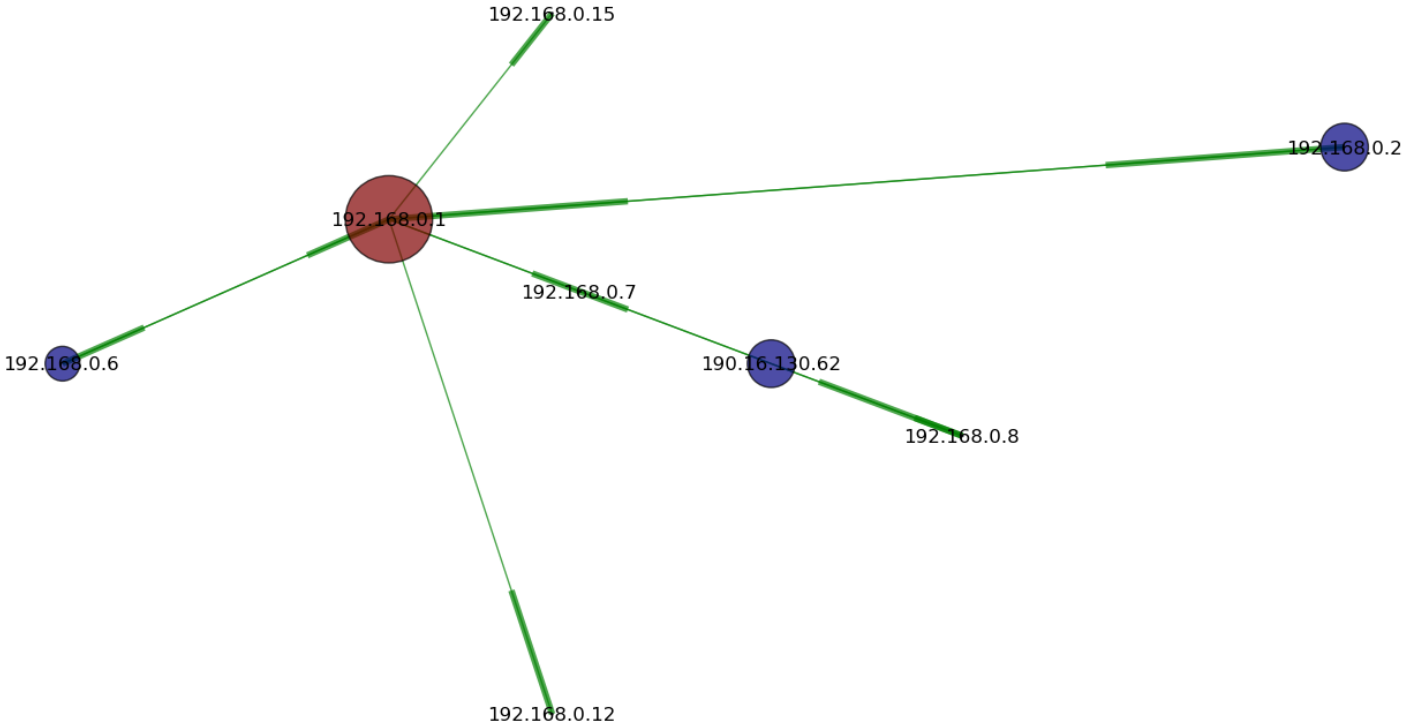

Figura 3: Red WiFi casa doméstica 1

En este gráfico, podemos identificar claramente a un nodo con gran actividad dentro de la red, coherentemente con los gráficos anteriores, es nuevamente en este gráfico el nodo 192.168.0.1 el de mayor actividad dentro de la red, correspondiéndose esta dirección con la del Router.

3.2. Red WiFi casa particular 2

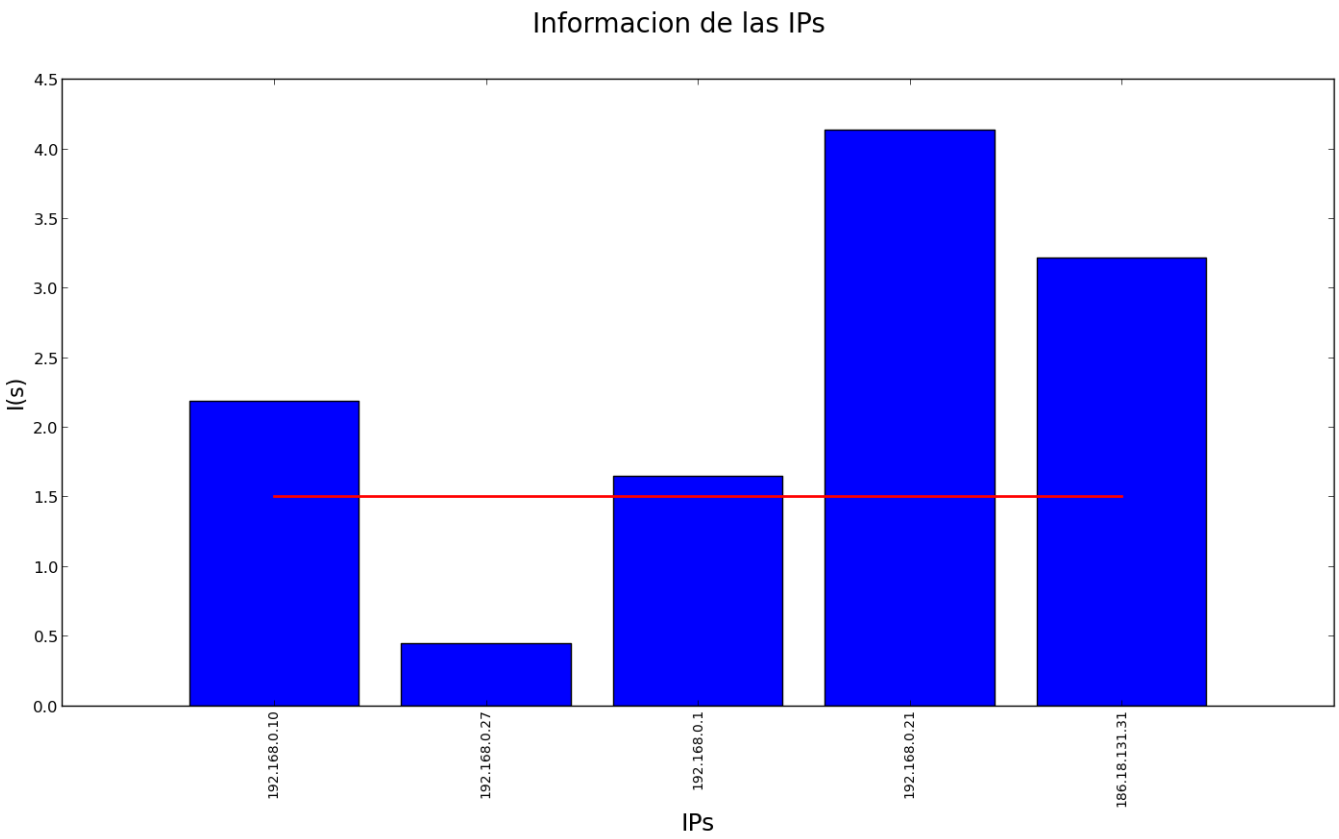

Figura 4: Fuente de información: IPs que envían

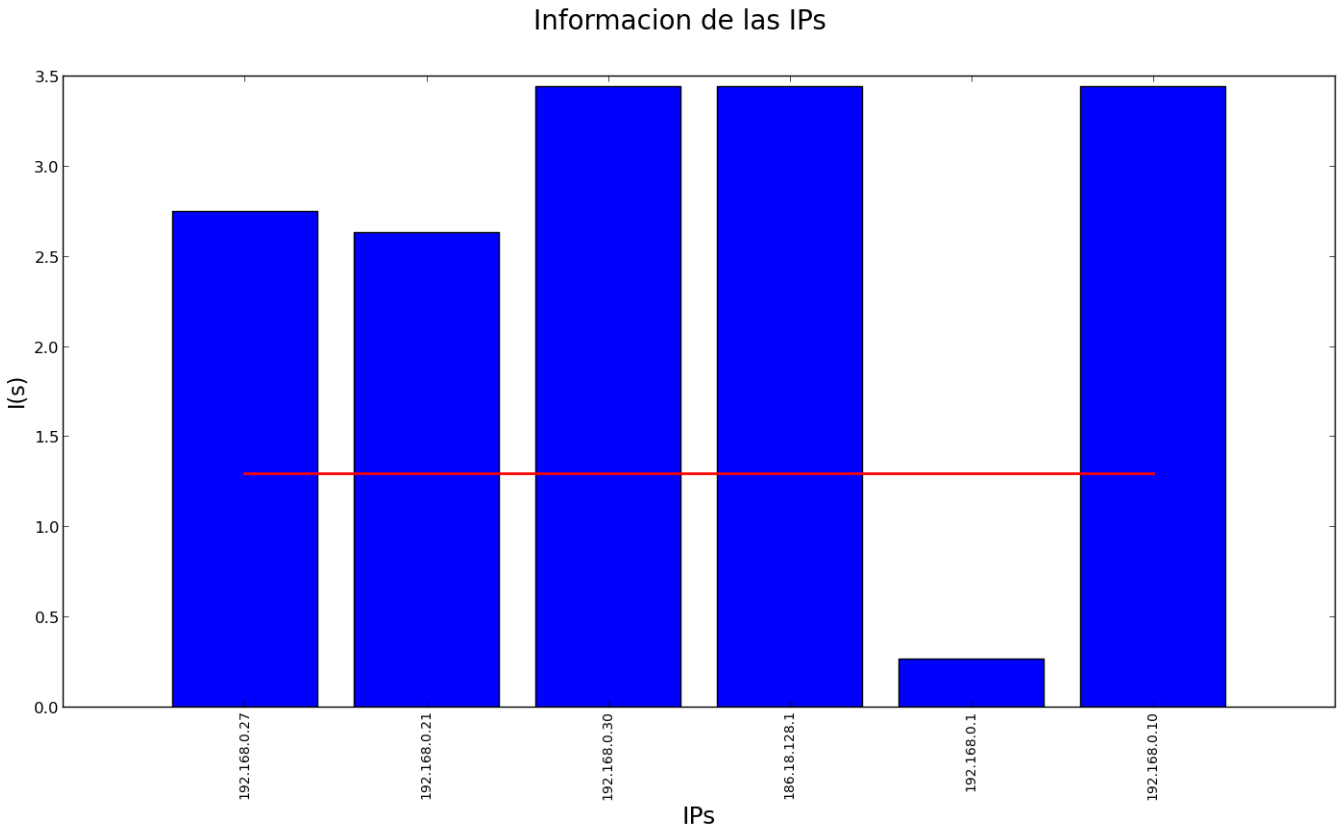

Figura 5: Fuente de información: IPs que reciben

Estos gráficos se corresponden con otra red de una casa. Sin embargo, nos devuelve algunos resultados que no estábamos esperando. En el primer gráfico nos encontramos con un evento distinguido, el cual se corresponde con la IP 192.168.0.27, sin embargo, si miramos el segundo gráfico el evento distinguido se corresponde con otra dirección IP, la 192.168.0.1. Esto nos plantea un interrogante sobre cuál se corresponde al Router. Nuestros conocimientos previos nos llevan a pensar que la segunda dirección se corresponde con el Router, y nos encargamos de chequearlo. La duda paso a ser, a qué host pertenecía la otra IP. Nos encontramos con que esta IP se corresponde al Servidor Web, Apache, utilizado localmente.

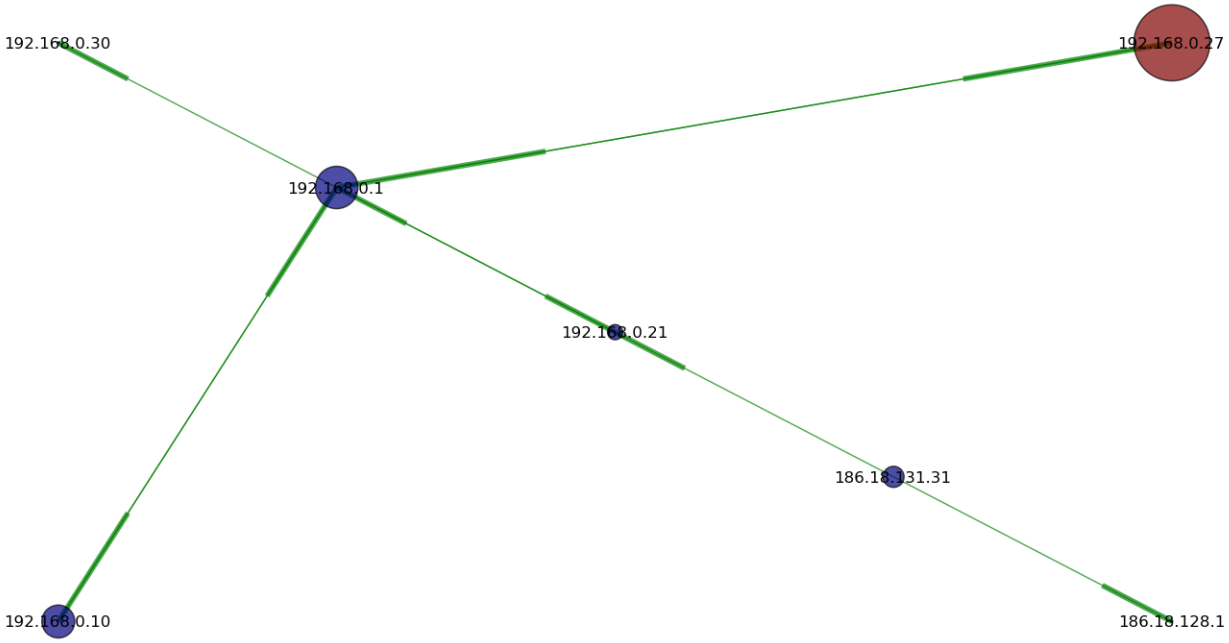

Figura 6: Red WiFi casa doméstica 2

Nuevamente, y coherentemente con lo mostrado en los gráficos anteriores, nos encontramos con un nodo significativo, pero que no se corresponde con el Router, sino con el Servidor Web. Sin embargo, debido a la cantidad de aristas con las que se conecta uno de los nodos, nos permite pensar que ese es el Router y lo comprobamos ya que su IP es 192.168.0.1

3.3. Red Ethernet Empresa (Recursiva)

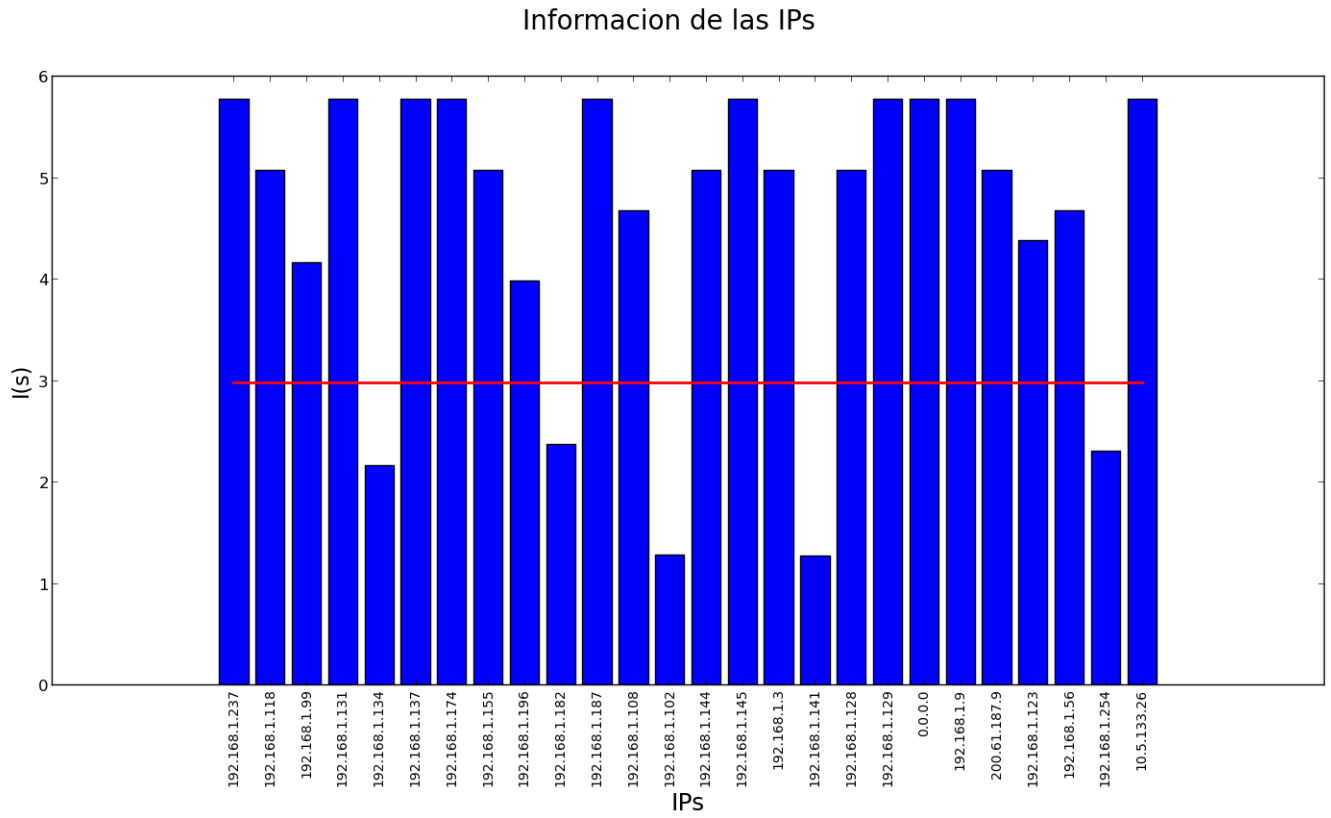

Figura 7: Fuente de información: IPs que envían

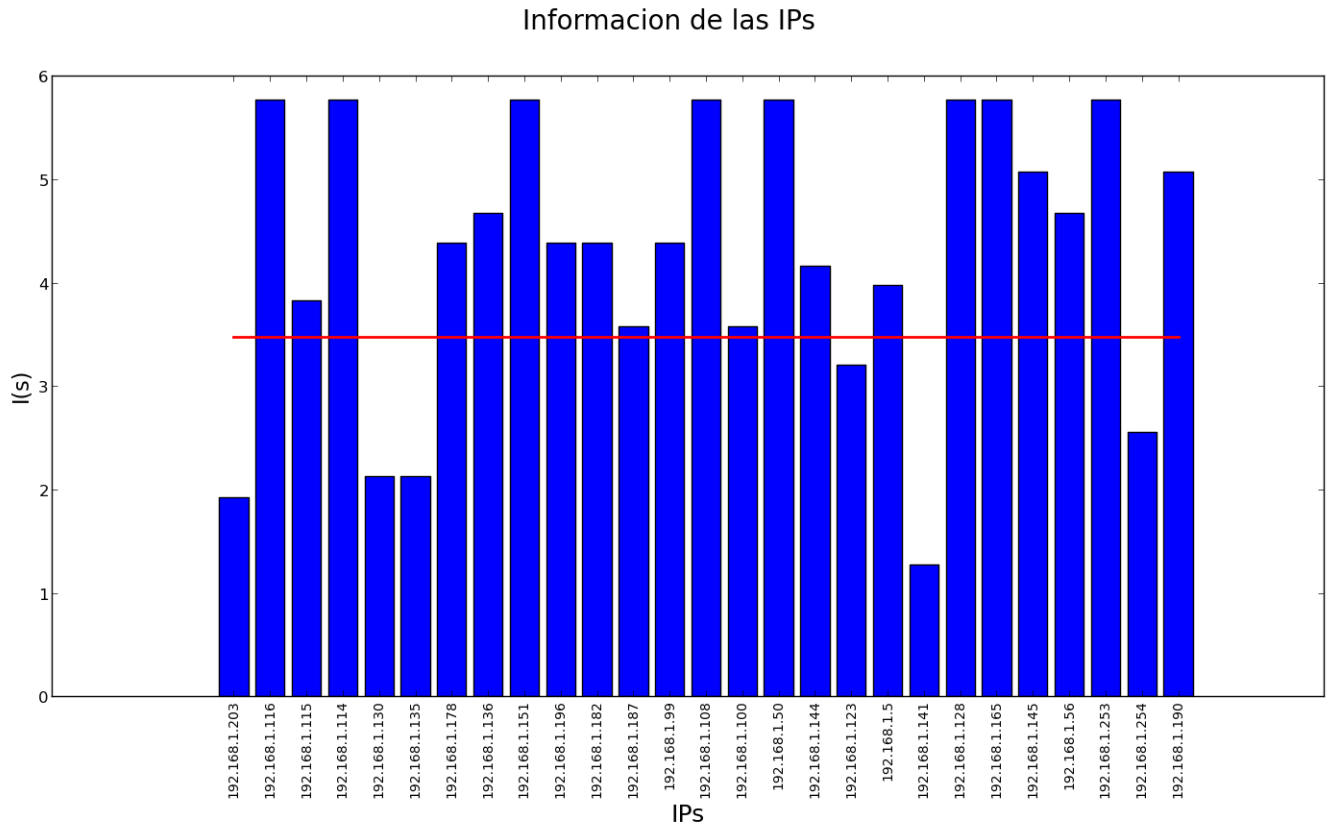

Figura 8: Fuente de información: IPs que reciben

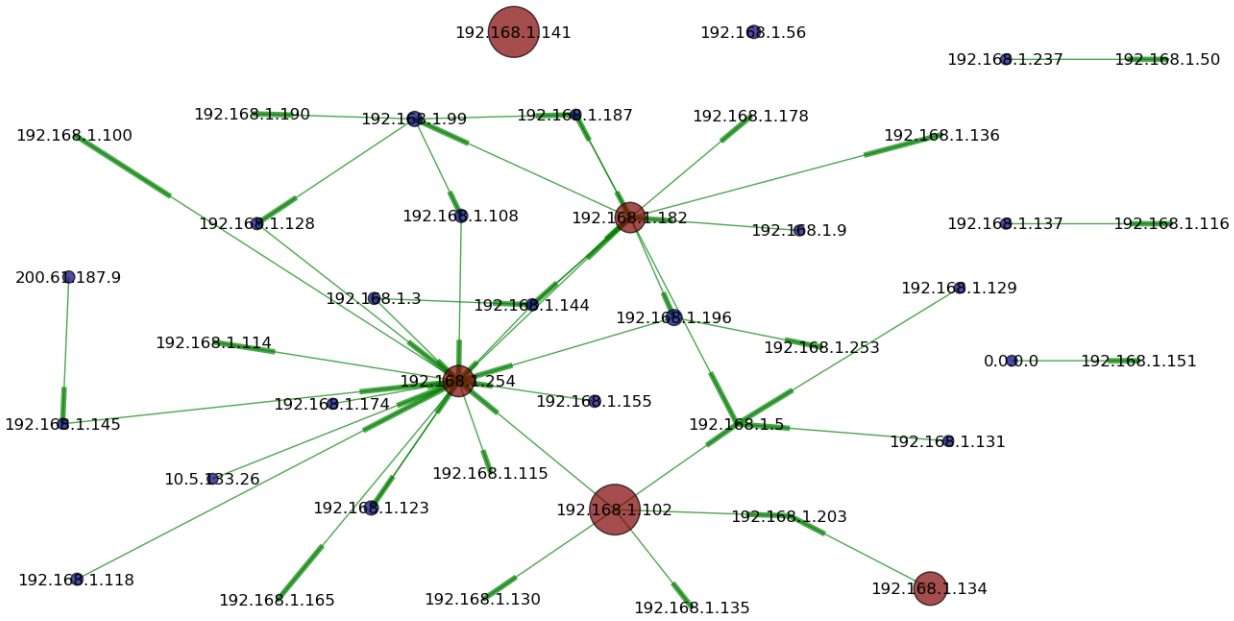

Figura 9: Red Ethernet de Recursiva

3.4. Red Ethernet Organismo (ORSNA)

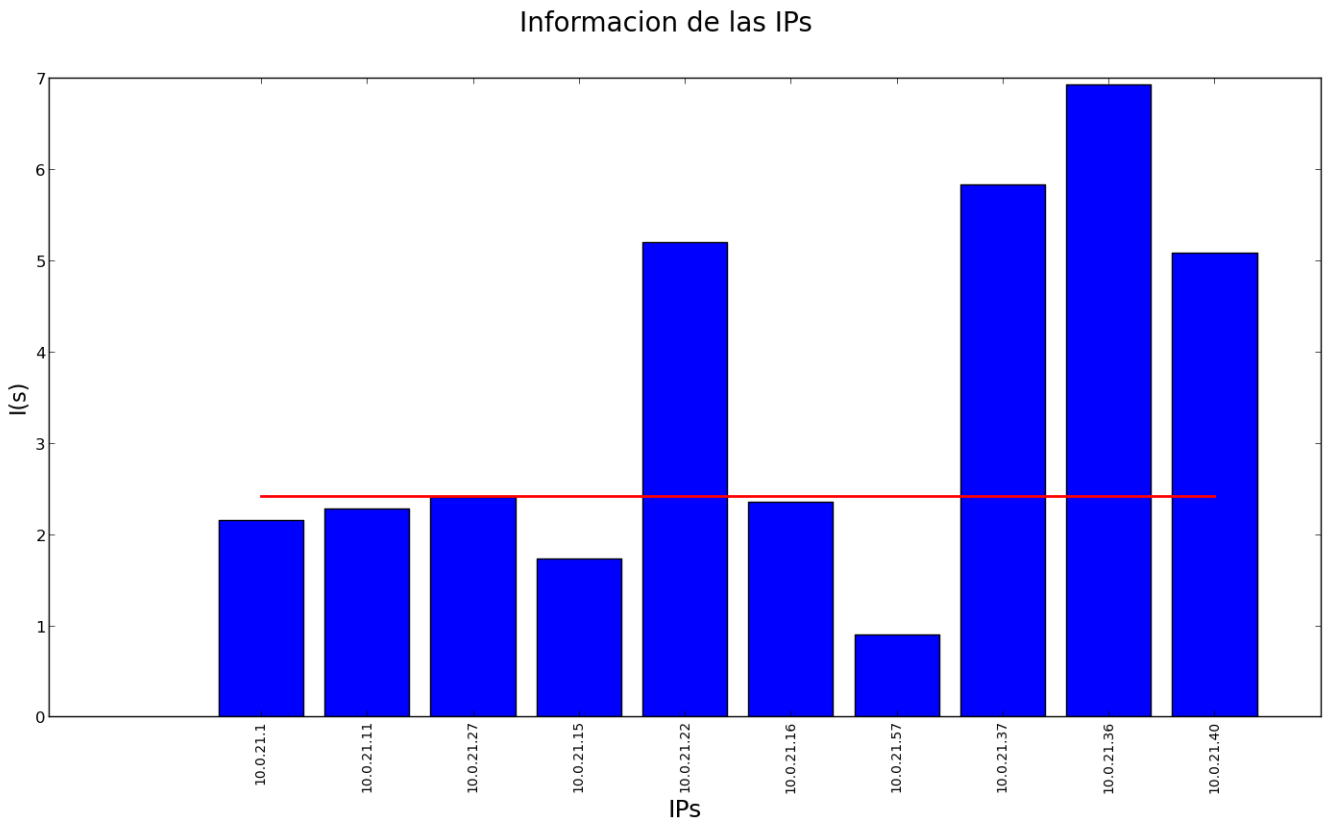

Figura 10: Fuente de información: IPs que envían

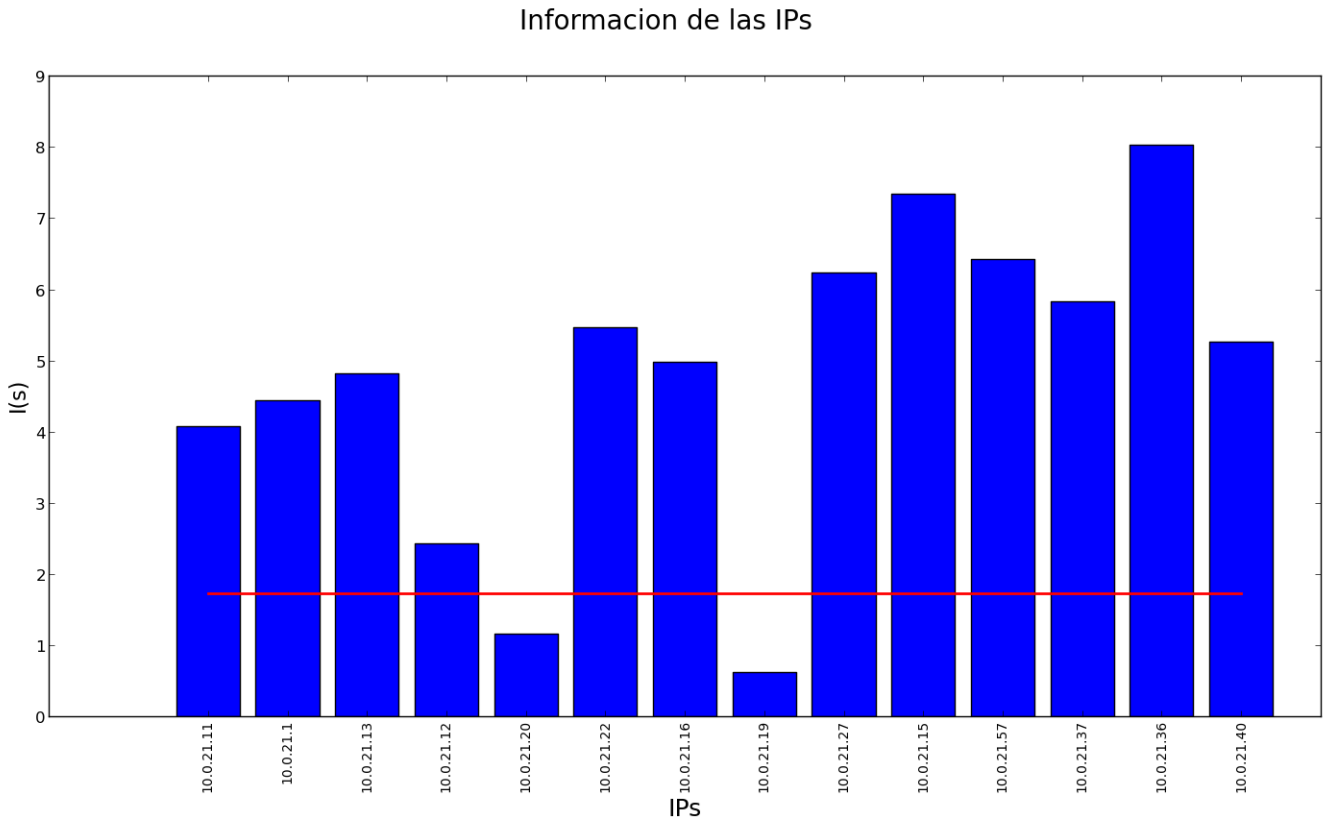

Figura 11: Fuente de información: IPs que reciben

A diferencia de las LAN anteriores, podemos notar que no está tan explícita la IP del router. Con ambas fuentes de información encontramos distintos hosts distinguibles. Rastreando a dónde pertenece cada dirección, encontramos que la IP 10.0.21.19 pertenecía a una impresora que dejó de funcionar y fue sacada del organismo, pero no fue sacada de la configuración de dispositivos de algunas computadoras, que son aquellas que realizan llamados who-has constantemente a esa dirección buscando la impresora perdida. Es por ello que esta dirección está por debajo de la entropía en el gráfico 11, debido a la cantidad de información que lleva este evento.

Además, tampoco se pudo rastrear a qué había pertenecido la dirección 10.0.21.20, otra dirección distinguible en los destinos, dado que actualmente no está asignada a ningún hardware. Esta dirección fue requerida constantemente por la IP 10.0.20.57 (al rededor de mil veces en 30 minutos). Esto es lo que hace que sea una fuente distinguida, pero no podemos deducir por qué, dado que nos falta información.

Con todos los datos sucios debido a los inconvenientes mencionados anteriormente, no se puede deducir fácilmente cuál es el router de la red, aunque rastreándolo pudimos comprobar que este se encuentra en la dirección 10.0.21.1.

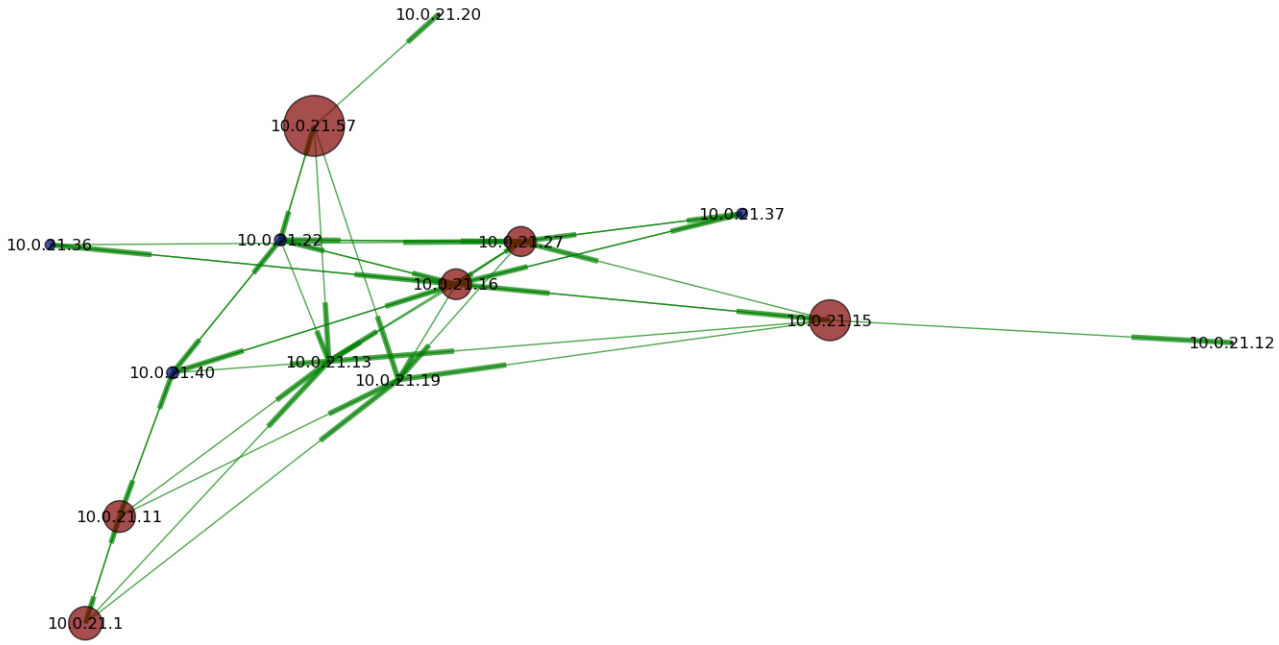

Figura 12: Red Ethernet de ORSNA

En este gráfico podemos ver nuevamente lo que mencionábamos antes. Las dirección IP 10.0.21.57 es la que más paquetes who-has mandó a la IP fantasma (10.0.21.20), y luego todas se mantienen más o menos igual debido que todas mandaron paquetes a la IP 10.0.21.19 que pertenecía a la impresora.

3.5. Red WiFi local comercial (McDonalds)

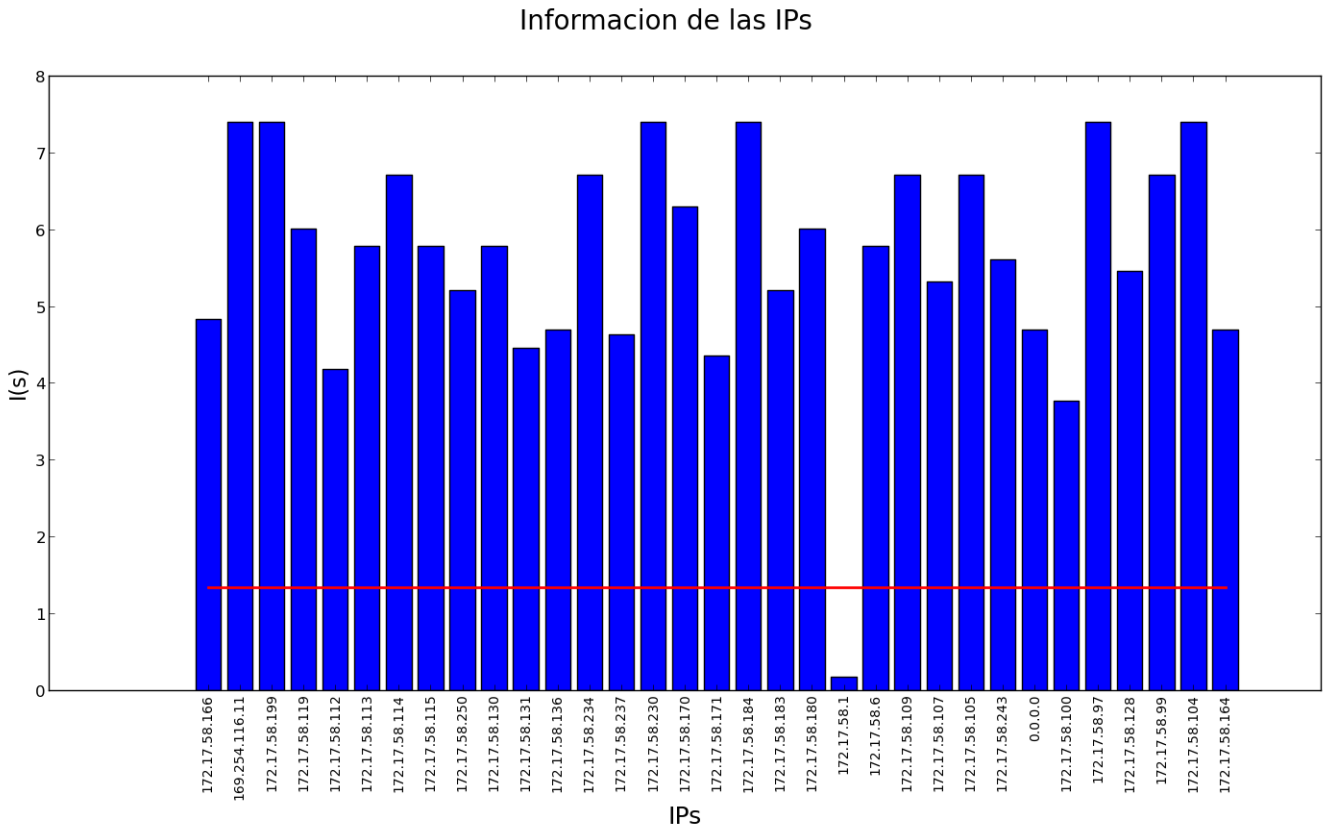

Figura 13: Fuente de información: IPs que envían

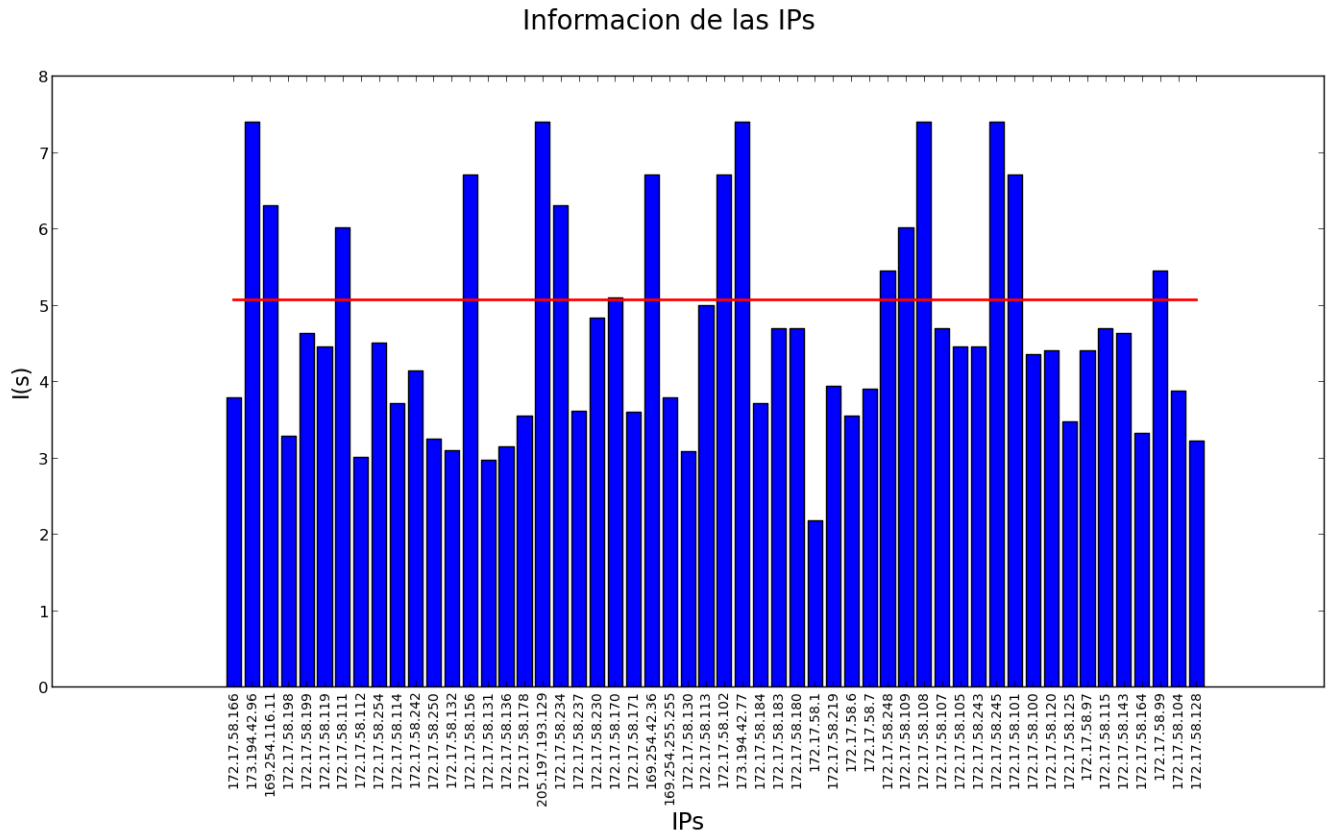

Figura 14: Fuente de información: IPs que reciben

Por último, pusimos a correr nuestro programa en la red WiFi del local comercial McDondalds y al analizarlo se ve claramente un solo nodo distinguido en el gráfico 13 al que podemos asumir sin dudar como el router. En el lapso de 40 minutos, el host 172.17.58.1 mandó 1372 ARP requests en contraste con los otros hosts que mandaron como mucho 38.

La otra fuente no aporta muchos datos ya que todos los nodos reciben muchos paquetes por igual.

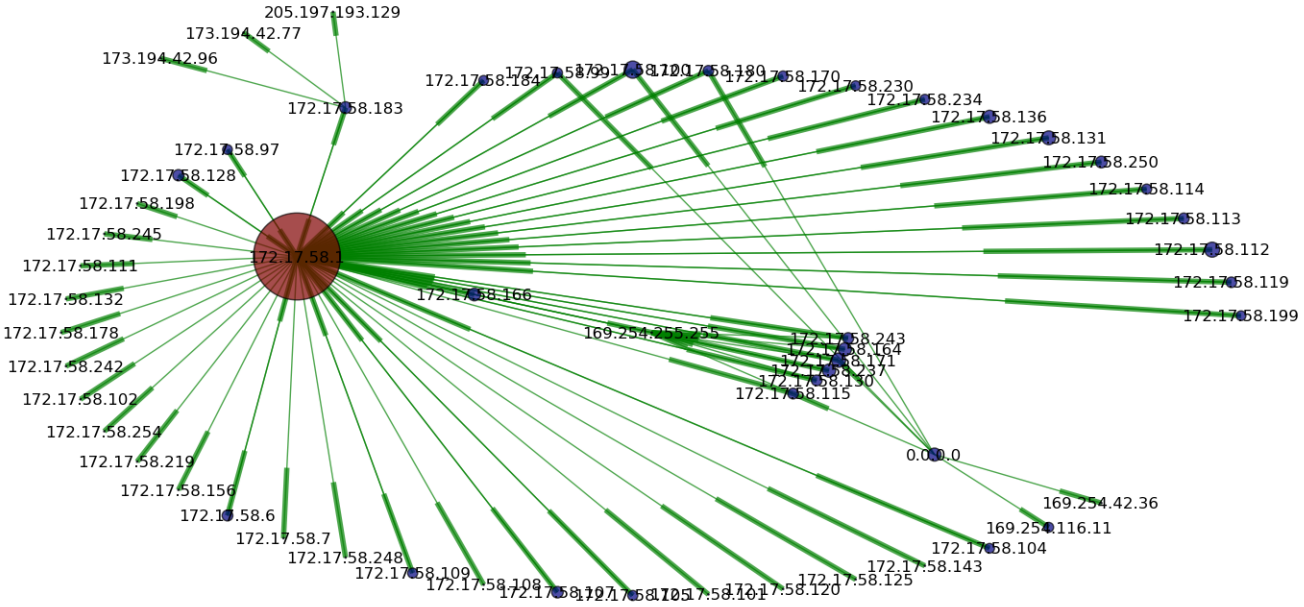

Figura 15: Red WiFi de McDonalds

4. Conclusiones

Cuando presentamos el trabajo práctico, marcabamos la influencia que tiene el rol de cada dispositivo con respecto al uso del protocolo ARP. Así, esperabamos que el Router sea un nodo distinguido en las redes que ibamos a analizar. En los distintos casos, nos encontramos con diferentes resultados. Algunos de ellos se correspondían con lo que esperábamos, pero en otros nos encontramos con algunos resultados inesperados.

Pudimos ver, que si bien el Router participa en una cantidad importante de los paquetes ARP que circulan en la red, podemos encontrarnos con distintos Hosts que cobran importancia. Sin embargo, dejando de lado el análisis meramente numérico, se puede observar que los nodos que representan a los Routers, en la gran mayoria de los casos, se encuentran rodeados por varios nodos, y presentan muchas aristas, con respecto a los demás.

Sin lugar a dudas, la gran variedad de redes, las distintas formas de configurarlas y la diversidad de Hosts que podemos encontrarnos, nos muestran que no existe una regla que nos diga, que rol cumple cada nodo. En las casas particulares y lugares públicos es efectivo usar la fuente de información que usa el evento de ip origen, porque hay un solo gateway y los demás hosts son en general terminales de usuarios normales como pc, notebooks o celulares que no interfieren en el análisis. A diferencia de una red en un lugar de trabajo, donde podemos encontrar distintos routers, y terminales de administradores que pueden estar haciendo mantenimiento o su propio estudio de la red, generando tráfico ARP para encontrar hosts caídos, por ejemplo, y que influye en los resultados finales, como ya vimos.

A pesar de esto, lo que si podemos encontrar es una generalidad, y ver que nuestra hipótesis primigenia de que el Router sea quien más utiliza el protocolo ARP es acertada, y en la mayoría de los escenarios es efectivo.