DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico N° 1

Teoría de las Comunicaciones
Segundo Cuatrimestre del 2016

Integrante	LU	Correo electrónico
Thibeault, Gabriel	114/13	gabriel.eric.thibeault@gmail.com
Guerson, Matias	925/10	matias.guerson@gmail.com

Reservado para la cátedra

Instancia	Docente	Nota
Primera entrega		
Segunda entrega		

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

1. Introducción

1.1. Información

La **información** de un evento aleatorio e con probabilidad $P(e)$ está dada por:

$$I(e) := -log_b(P(e))$$

para una cierta base b. La elección de dicha base determina la unidad de la información; en este trabajo nos limitaremos a base 2, por lo que la unidad que manejaremos son los bits.

La **entropía** de una variable aleatoria A es la esperanza de la información de A, y está dada por:

$$H(A) := \sum_{a \in A} P(a) * I(a) = -\sum_{a \in A} P(a) * log(P(a))$$

La entropía máxima de una variable aleatoria se da cuando los eventos son equiprobables. En particular, para una variable Bernoulli[1] equiprobable, la máxima entropía es de 1 bit.

Una **fuente** emite mensajes con ciertas probabilidades. Una **fuente de memoria nula** es una en la cual la probabilidad de cada mensaje no depende de los mensajes previos; viendo cada mensaje como una variable aleatoria, esto equivale a que sean independientes. Adicionalmente, si la probabilidad de cada mensaje es constante en el tiempo[2], estas variables además son idénticamente distribuidas.

La entropía de una fuente de memoria nula es la entropía de cada mensaje, que equivale a la información esperada de cada mensaje.

1.2. Capa de enlace y ARP

En la mayoría de los protocolos de la capa de enlace se emplean identificadores únicos; en el caso de Ethernet (802.3) y WLAN (802.11), este identificador lleva el nombre de MAC (*Media Access Control*) address.

Un frame de una red Ethernet puede ser transmitido de forma *unicast*, es decir para sólo un receptor[3], especificando como destino su MAC address; o de forma *broadcast*, es decir para todos los dispositivos de la red, marcando como desinto la MAC address FF:FF:FF:FF:FF:FF.

Generalmente, la transmisión *broadcast* se emplea para protocolos de control (por ejemplo ARP o detección de colisiones en Ethernet), mientras que la *unicast* podrá enviar datos[4].

También es ubicuo el uso de los identificadores únicos en la capa de red; para el protocolo más común (IP), se utilizan Direcciones IP.

Para relacionar un identificador de capa de red con uno de capa de enlace[5], se emplea el protocolo ARP (*Address Resolution Protocol*). En la figura 1 se puede ver la estructura de un paquete ARP.

Para encontrar la MAC address correspondiente a una Dirección IP conocida, un dispositivo envía un paquete a la red de forma *broadcast*. Marca en 1 el campo Operation (*who-has*), anota su Dirección IP y su MAC address en los campos *Sender's Protocol Address* y *Sender's Hardware Address*, respectivamente, y escribe la Dirección IP deseada en el field *Target Protocol Address*.

Cuando un dispositivo recibe el paquete e identifica a su propia Dirección IP como la *Target Protocol Address*, responde enviando otro paquete ARP al emisor original. En este caso, el campo Operation se settea en 2 (*is-at*). Ambos campos de *Sender's Address* nuevamente se completan con sus direcciones, mientras que los de *Target Address*, con las direcciones provistas por el emisor en el paquete *who-has* original.

Cabe destacar que un paquete *who-has* debe ser enviado de forma *broadcast* (ya que la MAC address del receptor es desconocida), mientras que uno *is-at* se transmite de forma *unicast*, pues el emisor original envió su MAC address en el request original.

Adicionalmente, hay dos casos de uso especiales de ARP: *ARP probing* y *gratuitous ARP*. El primero se emplea para evitar colisiones en los identificadores de IP (específicamente en IPv4, la versión más común actualmente), y se destaca al señalar el campo *Sender's Protocol Address* con todos 0[6]. El segundo se usa como anuncio, y en éste se marcan los campos de *Sender's Protocol Address* y *Target Protocol Address* con la dirección IP del dispositivo que realiza el anuncio.

2. Desarrollo

2.1. Fuente S

Definiremos una fuente de memoria nula S en base a los frames de capa de enlace capturados. La fuente consiste en dos mensajes: un frame fue transmitido de forma *broadcast*, o éste fue transmitido de forma *unicast*.

[1] Es decir, una que admite sólo dos eventos posibles.

[2] A partir de ahora, asumiremos que lo es.

[3] Debido a la naturaleza de Ethernet, se transmite a toda la red, pero los otros receptores normalmente descartarán los frames que no les son destinados.

[4] También podrá ser utilizada por algunos protocolos de control, por ejemplo ARP, como se detallará a continuación.

[5] En nuestro caso, Direcciones IP y MAC addresses, respecitvamente. De aquí en adelante utilizaremos estos términos al referirnos a ARP.

[6] El valor específico es 0.0.0.0.

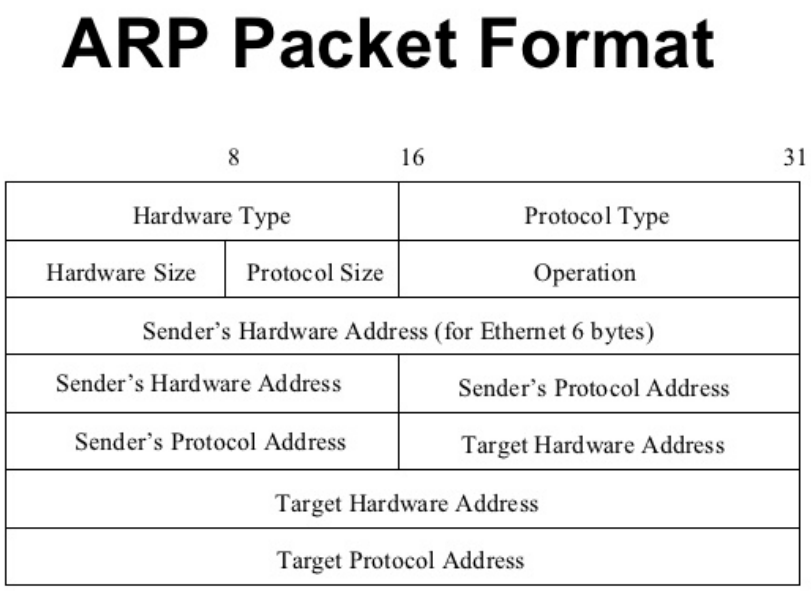

Figura 1: Estructura de un paquete ARP.

2.2. Elección de la fuente S_1

Definiremos una fuente de memoria nula S_1 en base a las Direcciones IP de los paquetes ARP. Deberemos tomar diversas decisiones para definirla correctamente para poder distinguir los nodos apropiados.

En primer lugar: debemos elegir si tomar los paquetes *who-has*, *is-at*, o ambos. En la mayoría de los casos, un *who-has* será respondido por exactamente un *is-at* correspondiente, a menos que el receptor deseado no pueda recibir el paquete o emitir la respuesta, o que haya dos dispositivos con una misma MAC address que intenten responder a la vez. Por ende, la información del *is-at* será redundante con la del *who-has*, a menos que se produzca un error (lo que, de tomar ambos, agregaría errores a las mediciones).

Consecuentemente, tomaremos sólo uno. Ya que el *who-has* se transmite de forma *broadcast*, mientras que el *is-at*, de forma *unicast*[7], tomaremos el primero.

En segundo lugar, debemos decidir si emplear el origen del *who-has*, su destino, o ambos como el mensaje de la fuente. Esta decisión no la tomaremos de antemano, sino que observaremos los grafos resultantes de los experimentos y en base a ellos decidiremos cuál es la opción más acertada.

En último lugar, debemos decidir si permitir mensajes repetidos[8]. Si bien esto no es ilógico desde el punto de vista del modelo de fuente de memoria nula planteado, los paquetes ARP repetidos no deberían ser necesarios: una vez que se envía un *who-has* por una cierta dirección IP y éste es respondido por un *is-at*, la relación entre esta dirección y la MAC address provista debería persistirse en una tabla del emisor; paquetes repetidos podrían ser síntomas de que el *who-has* original no tuvo respuesta, por lo que otros posteriores fueron requeridos.

Creemos que por esta razón no deberíamos considerar paquetes repetidos, pero de todas formas juzgaremos ambos procedimientos en base a los resultados de los experimentos.

Descartaremos paquetes correspondientes a *gratuitous ARPs* y *ARP probings*, ya que sus características son anómalas y no nos ayudarán a distinguir nodos.

Definimos a un nodo p como **distinguido** si:

$$I(p) < H(S_1)$$

Los grafos que emplearemos para representar la red subyacente de mensajes ARP serán independien-

[7] Si bien realizaremos las mediciones en modo promiscuo, la presencia de *switches* puede evitar que veamos este tipo de paquetes si no están destinados a nuestro dispositivo, lo que generaría aún más errores en las mediciones.

[8] Es decir, si considerar repetidas veces múltiples paquetes ARP con igual origen y destino.

tes de las elecciones que tomemos respecto de la fuente. En particular, éstos consistirán en digrafos con loops[9], donde hay un eje de un nodo a otro si el primero emite un *who-has* preguntando por la Dirección IP del segundo.

2.3. Experimento 1: red inalámbrica de los laboratorios del DC

Para este experimento evaluamos la red inalámbrica de los laboratorios del DC.

2.3.1. Fuente S

A continuación podemos ver la fuente S propuesta, modelada con los resultados del experimento:

Mensaje	Probabilidad	Información [bits]
Unicast	0.773	0.371
Broadcast	0.227	2.141

Entropía de la fuente: 0.772 bits. Entropía máxima: 1 bit.

Observamos que la entropía de la fuente es menor que la máxima, ya que las transmisiones *unicast* son casi 3 veces más probables que las *broadcast*. Esto nos provee una cota inferior para el *overhead* impuesto por los protocolos de control: al menos 22.7 % de los frames no transmiten datos.

2.3.2. Estructura de la red en base a los paquetes ARP

En las figuras 2 y 3 se pueden ver los grafos[10] de la red subyacente de mensajes ARP.

Las direcciones IP de la red son de la forma 10.2.X.Y, con una sola excepción que mencionaremos más adelante. Éstas son direcciones IP privadas[11].

La red se presenta altamente fragmentada; el grafo posee múltiples componentes conexas. Adicionalmente, vemos repetido un patrón entre varias de estas componentes: un nodo central, con una dirección IP de la forma 10.2.X.254 o 10.2.X.249, que envía paquetes a múltiples hojas[12]. Esta estructura es consistente con el comportamiento esperado de Default Gateways.

Hay un nodo claramente destacado en la red, el de dirección IP 10.2.203.254, que tanto envía como recibe paquetes de un gran número de hojas.

Se advierten diversas anomalías: en primer lugar, loops en el grafo, que como mencionamos previamente, se deben a *gratuitous ARPs*; en segundo lugar, la dirección 0.0.0.0 se hace presente en el grafo, siempre como origen, lo que ejemplifica *ARP probing*; finalmente, una única IP que no comienza con 10.2, la 169.254.255.255. El rango 169.254.0.0-169.254.255.255 está reservado; se asigna cuando un dispositivo no tiene IP estática, y el protocolo dinámico[13] utilizado falla.

Mensaje	Información [bits]	Distinguido?
10.2.3.249	7.58	No
10.2.202.249	7.58	No
10.2.6.249	6.58	No
10.2.7.254	6.00	No
10.2.0.254	6.00	No
10.2.0.249	6.00	No
10.2.6.254	5.26	Sí
10.2.1.254	4.78	Sí
10.2.1.249	4.78	Sí
10.2.3.254	4.78	Sí
10.2.2.254	4.78	Sí
10.2.2.249	4.58	Sí
10.2.4.254	4.58	Sí
10.2.5.254	4.58	Sí
10.2.7.249	4.26	Sí
10.2.203.254	2.94	Sí

Tabla 1: Información de los nodos de la fuente S_1 en el experimento 1, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP en el campo *Sender's Protocol Address* de un paquete ARP *who-has*.

2.3.3. Fuente S_1

En la tabla 1 podemos ver la información de ciertos[14] mensajes de la fuente S_1, sin paquetes repetidos y tomando sólo el *Sender's Protocol Address* de los paquetes.

La entropía de la fuente es de 5.61 bits, siendo la máxima 32 bits.

Vemos que no se producen falsos positivos: todos los nodos que la fuente distingue son los que destacamos previamente, con direcciones IP de la forma 10.2.X.249 o 10.2.X.254. Sin embargo, vemos ciertos potenciales falsos negativos.

Los tres primeros, es decir el 3.249, el 202.249 y el 6.249[15], si bien sus IPs son de la forma destacada, no parecen exhibir el comportamiento de los otros[16]

[9] Como mencionamos en la Introducción, en un *gratuitous ARP*, *Sender's Protocol Address* = *Target Protocol Address*. Para poder observar este fenómeno en el grafo, permitiremos loops.

[10] Ambos grafos representan la misma red. Sin embargo, el tamaño del grafo 2 puede dificultar un análisis detallado, por lo que en la figura 3 colapsamos ciertos nodos con iguales vecinos (que se muestran en rojo). Mantuvimos el grafo original ya que una mirada rápida ofrece más información concerniente a la topología de la red.

[11] Todo el rango 10.0.0.0-10.255.255.255 es privado.

[12] Se presenta una estructura de estrella, o cercana.

[13] DHCP siendo actualmente el más común.

[14] La alta cantidad de nodos dificulta seriamente la presentación de estos resultados, tanto en forma de gráficos como de tablas. En las siguientes tablas listaremos todos los nodos distinguidos, y los nodos que, en base a ciertos criterios, nos parecieron destacados pero la fuente no distinguió.

[15] Obviando de las direcciones el 10.2. inicial.

[16] En el grafo no se presentan en el centro de una estructura

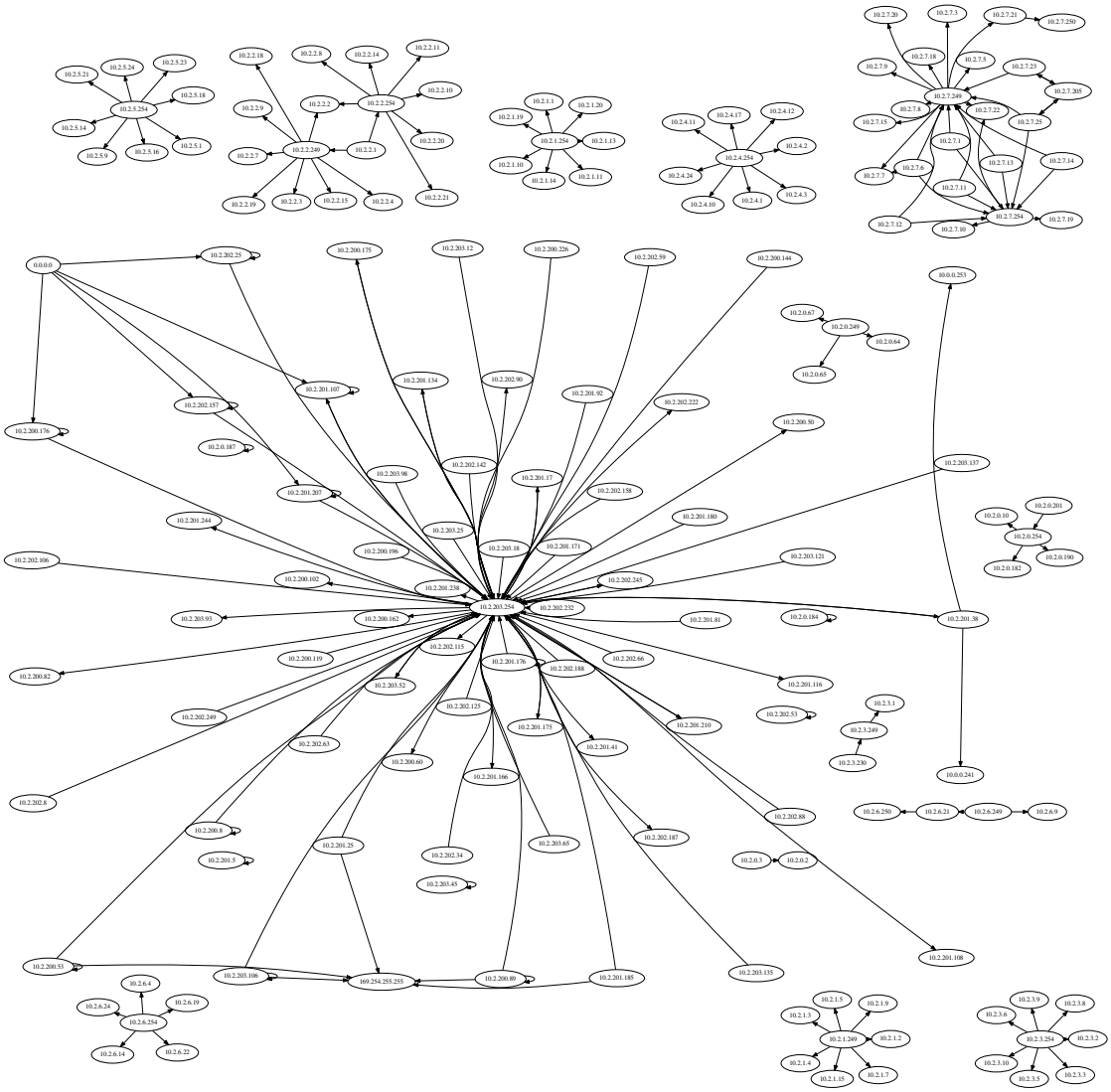

Figura 2: Grafo de la red subyacente de mensajes ARP en el experimento 1, sin colapsar nodos .

ni actuar como Default Gateways. Por ende, concluimos que no clasificarlos como distinguidos no es una falencia de la fuente.

Sin embargo los tres restantes, el 7.254, el 0.249 y el 0.254, presentan el comportamiento mencionado y exhiben la estructura de estrella, pero no son clasificados como distinguidos. En el caso del 0.249, esto se debe a que no tiene suficientes vecinos, por lo que su información es relativamente baja; es posible que esta falencia pueda ser subsanada al considerar paquetes repetidos. Por otro lado, el 0.254 y el 7.254, tienen más vecinos pero varios de éstos presentan ejes en el otro sentido que el aceptado por la fuente[17]; es posible que este error sea resuelto aceptando ejes en ambas direcciones.

En la tabla 2 podemos ver la información de ciertos mensajes de la fuente S_1, sin paquetes repetidos y tomando tanto el *Sender's Protocol Address* como el *Target Protocol Address* de los paquetes.

La entropía de la fuente es de 6.28 bits, siendo la máxima 32 bits. Cabe destacar que la entropía es mayor que la de la fuente previa. Muchos paquetes son enviados por un nodo destacado a direcciones que no aparecen nuevamente; la fuente previa sólo cuenta al nodo destacado, que al aparecer múltiples veces provee poca informacion, mientras que esta fuente cuenta también al nodo más raro, cuya información es más alta.

Vemos que uno de los falsos negativos que habíamos mencionado, el 7.254, es considerado distinguido en esta nueva fuente. Se presenta otro nuevo nodo como distinguido: el 169.254.255.255. Previamente habíamos mencionado que ésta era una IP reservada con características específicas; no es absurdo considerarla destacada. Sin embargo, si la intención es que

de estrella o similar.

[17] Es decir, este nodo es el destino de varios paquetes ARP, no el origen.

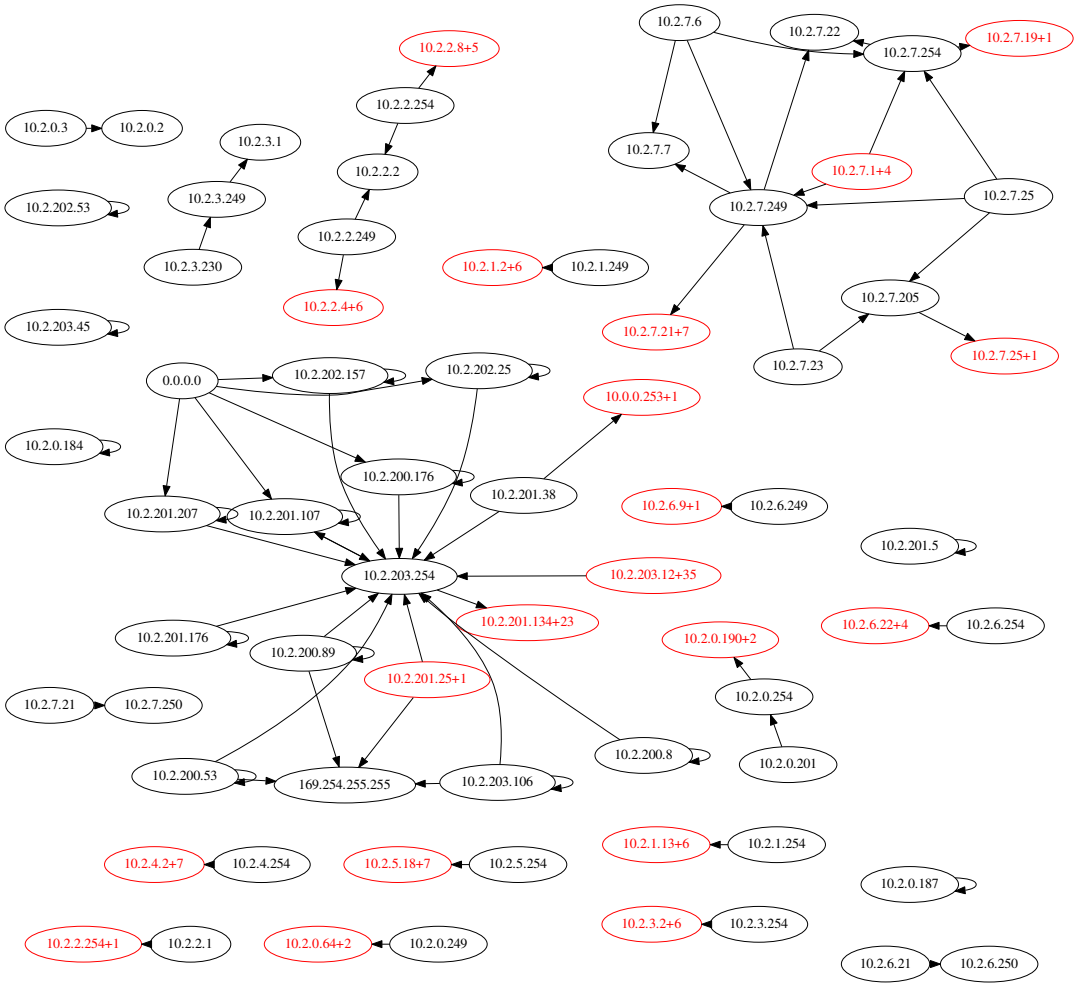

Figura 3: Grafo de la red subyacente de mensajes ARP en el experimento 1, colapsando nodos.

todo nodo destacado sea un Default Gateway, esto nos muestra que se deben eliminar ciertos rangos de direcciones particulares.

En la tabla 3 podemos ver la información de ciertos mensajes de la fuente S_1, con paquetes repetidos y tomando sólo el *Sender's Protocol Address* de los paquetes.

Entropía de la fuente: 5.20 bits. Entropía máxima: 32 bits.

Los resultados presentan una extrema cantidad de falsos positivos y negativos, por lo que concluimos que esta fuente no sirve para nuestros propósitos.

Debido a la mayor entropía y a la menor cantidad de falsos negativos de la fuente sin repetidos y con origen y destino, concluimos que ésta es preferible, al menos en las condiciones de este experimento.

2.4. Experimento 2: red de oficina de trabajo

Para este experimento realizamos las mediciones sobre una red WiFi laboral durante una hora.

2.4.1. Fuente S

Vemos a continuación las métricas de la fuente S propuesta, modelada con los resultados del experimento:

Mensaje	Probabilidad	Información [bits]
Unicast	0.464	1.109
Broadcast	0.536	0.898

Entropía de la fuente: 0.996 bits. Entropía máxima: 1 bit.

Podemos observar que los dos tipos de paquetes son casi equiprobables, por lo que la entropía es muy cercana a la máxima.

Un resultado llamativo es que más de la mitad de

Mensaje	Información [bits]	Distinguido?
10.2.202.249	8.55	No
10.2.3.249	7.55	No
10.2.6.249	7.55	No
10.2.0.249	6.96	No
10.2.0.254	6.55	No
10.2.6.254	6.22	Sí
169.254.255.255	6.22	Sí
10.2.3.254	5.74	Sí
10.2.1.249	5.74	Sí
10.2.1.254	5.74	Sí
10.2.5.254	5.55	Sí
10.2.4.254	5.55	Sí
10.2.2.254	5.55	Sí
10.2.2.249	5.38	Sí
10.2.7.254	5.22	Sí
10.2.7.249	4.38	Sí
10.2.203.254	2.34	Sí

Tabla 2: Información de los nodos de la fuente S_1 en el experimento 1, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP tanto en el campo *Sender's Protocol Address* como en el *Target Protocol Address* de un paquete ARP *who-has*.

los paquetes hayan sido envíados como *broadcast*. Por tal motivo decidimos analizar en mayor profundidad los paquetes enviados como *broadcast*.

En la figura 4 podemos ver los resultados de evaluar los diversos tipos de los paquetes *broadcast*.

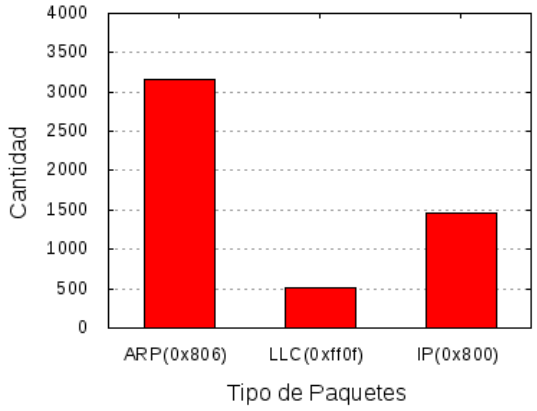

Figura 4: Tipos de paquetes broadcast en el experimento 2.

Como podemos observar en la figura 4, la mayor cantidad de paquetes son de tipo ARP. Podemos notar una pequeña cantidad de paquetes con tipo LLC (Logical Link Control). Éste es propio del estándar IEEE 802.2 que define el control de enlace lógico para redes de área local en el modelo OSI. Estos dos tipos de paquetes son utilizados por protocolos de control y su funcionamiento requiere transmisión *broadcast*. Lo llamativo aquí es la gran cantidad de paquetes *broadcast* de tipo IP.

Mensaje	Información [bits]	Distinguido?
10.2.3.249	9.63	No
10.2.202.249	9.63	No
10.2.6.249	8.63	No
10.2.7.254	7.63	No
10.2.1.254	6.63	No
10.2.0.254	6.63	No
10.2.6.254	6.63	No
10.2.2.254	6.46	No
10.2.2.249	6.17	No
10.2.5.254	6.04	No
10.2.1.249	5.93	No
10.2.4.254	5.82	No
10.2.7.249	5.82	No
10.2.3.254	5.54	No
10.2.200.144	4.72	Sí
10.2.200.176	4.68	Sí
10.2.7.12	4.46	Sí
10.2.7.25	4.42	Sí
10.2.7.14	4.38	Sí
10.2.2.1	4.27	Sí
10.2.7.11	4.24	Sí
10.2.203.254	4.07	Sí
10.2.7.6	4.07	Sí
10.2.7.1	4.07	Sí
10.2.7.13	3.99	Sí
10.2.0.249	3.82	Sí

Tabla 3: Información de los nodos de la fuente S_1 en el experimento 1, tomando paquetes repetidos y considerando como mensaje la ocurrencia de una IP en el campo *Sender's Protocol Address* de un paquete ARP *who-has*.

Para mejor comprender esto, graficamos la cantidad de paquetes de este tipo enviados por IP de origen, como se puede ver en la figura 5.

Observemos los tres primeros *sources* que sobresalen del resto en cuanto a cantidad de paquetes enviados. Éstos son los nodos cuyas MAC addresses son: 00:10:75:2d:f4:67 (386 paquetes), c4:85:08:2f:0f:e4 (294 paquetes) y 00:01:02:6c:95:05 (157 paquetes). Analizando los paquetes de tipo IP que provenían de estos nodos en modo *broadcast*, encontramos que emitían los siguientes mensajes:

- 00:10:75:2d:f4:67 : Hello there. I am at 192.168.1.120. Time is 1474050708 and I am hungry.Hostname: backupdyd.seagateshare.com Notamos que se trata de un software de backup que podría estar notificando a todos los nodos sus datos para posteriores procesos.
- c4:85:08:2f:0f:e4 : De este *source* no pudimos obtener mucha información como para poder determinar el propósito de los paquetes *broadcast*.
- 00:01:02:6c:95:0: 9016 3 ipp://192.168.1.150:631/printers/HP-LaserJet-

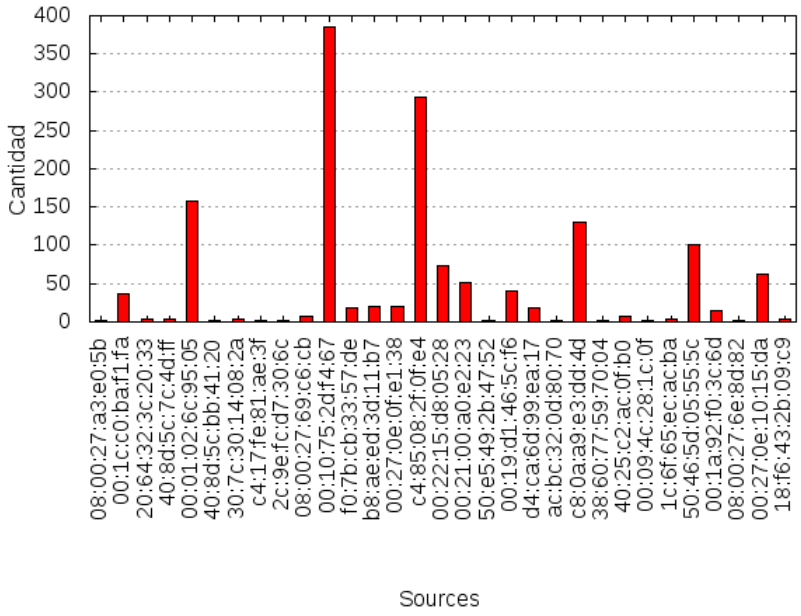

Figura 5: MAC addresses de los *sources* de paquetes IP broadcast.

P1006 "HP LaserJet P1006HP LaserJet P1006 Foomatic/foo2xqx (recommended)"job-shee ts=none,none lease-duration=300
Se trata de mensajes emitidos por impresoras que eventualmente podrían querer informar sus status a todos los nodos.

Por lo tanto, concluimos que se trata de paquetes de control, no de la red en sí misma, sino de protocolos propios de ciertos nodos.

2.4.2. Estructura de la red en base a los paquetes ARP

En la figura 6 se puede ver el grafo de la red subyacente de mensajes ARP.

Las direcciones IP de la red son de la forma 192.168.1.X, con ciertas excepciones: se observa la dirección 0.0.0.0 y la 169.254.255.255, ambas detalladas en el experimento previo; vemos dos nodos con las direcciones 192.9.200.100 y 192.9.200.1. Adicionalmente, éstos son los únicos dos nodos que no están conectados al resto de la red. Este comportamiento es anómalo, por lo que creemos que son IPs con un uso particular, específico a la red, ya que no parecen ser reservadas.

En el grafo se destaca claramente un nodo, el 192.168.1.1, cuyo comportamiento es consistente con lo esperado de un router[18]. Luego vemos un grupo de vertices con un muy alto grado de entrada y salida; sin embargo, no parecen actuar como Default Gateways.

[18]Es común que los routers ocupen la última o primera, como en este caso, dirección de la red.

Mensaje	Información [bits]	Distinguido?
192.168.1.1	4.86	No
192.168.1.194	4.50	Sí
192.168.1.127	4.50	Sí
192.168.1.90	4.34	Sí
192.168.1.99	4.34	Sí
192.168.1.121	4.34	Sí
192.168.1.158	4.21	Sí
192.168.1.205	4.08	Sí
192.168.1.117	3.34	Sí
192.168.1.235	2.96	Sí

Tabla 4: Información de los nodos de la fuente S_1 en el experimento 2, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP en el campo *Sender's Protocol Address* de un paquete ARP *who-has*.

2.4.3. Fuente S_1

En la tabla 4 podemos ver la información de ciertos mensajes de la fuente S_1, sin paquetes repetidos y tomando sólo el *Sender's Protocol Address* de los paquetes.

La entropía de la fuente es de 4.61 bits, siendo la máxima 32 bits.

El resultado es el opuesto al esperado: los nodos distinguidos son los pertenecientes al grupo de vértices mencionado con alto grado de interconexión pero sin comportamiento de router, mientras que el nodo efectivamente identificado como Default Gateway no es distinguido.

En la tabla 5 podemos ver la información de ciertos mensajes de la fuente S_1, sin paquetes repetidos y tomando tanto el *Sender's Protocol Address* como el

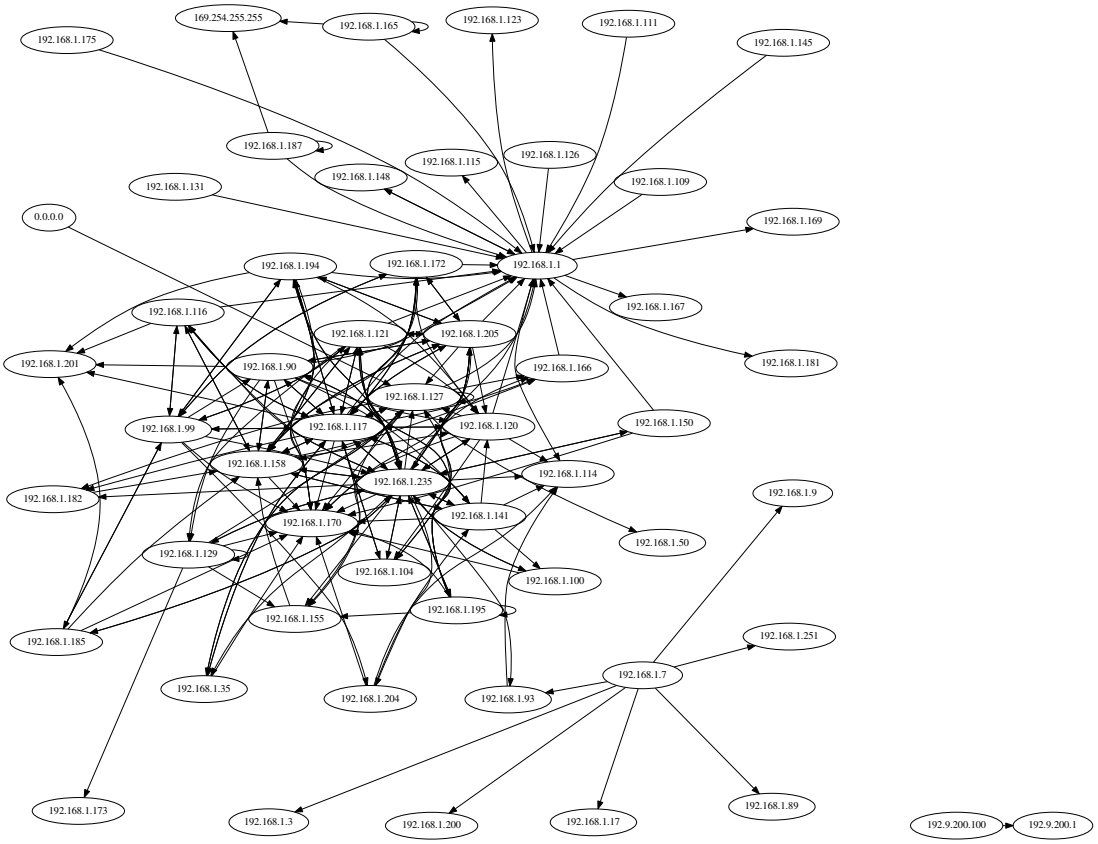

Figura 6: Grafo de la red subyacente de mensajes ARP en el experimento 2.

Mensaje	Información [bits]	Distinguido?
192.168.1.194	4.86	Sí
192.168.1.90	4.76	Sí
192.168.1.99	4.67	Sí
192.168.1.120	4.67	Sí
192.168.1.121	4.67	Sí
192.168.1.127	4.67	Sí
192.168.1.170	4.50	Sí
192.168.1.205	4.34	Sí
192.168.1.158	4.08	Sí
192.168.1.1	3.86	Sí
192.168.1.117	3.58	Sí
192.168.1.235	3.24	Sí

Tabla 5: Información de los nodos de la fuente S_1 en el experimento 2, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP tanto en el campo *Sender's Protocol Address* como en el *Target Protocol Address* de un paquete ARP *who-has*.

Target Protocol Address de los paquetes.

La entropía de la fuente es de 4.93 bits, siendo la máxima 32 bits.

Nuevamente se presentan los falsos positivos, pero el nodo que identificamos como router se considera ahora distinguido.

2.5. Experimento 3: red doméstica

Realizamos este experimento sobre una red doméstica.

2.5.1. Fuente S

A continuación podemos ver la fuente S propuesta, modelada con los resultados del experimento:

Mensaje	Probabilidad	Información [bits]
Unicast	0.848	0.237
Broadcast	0.152	2.720

Entropía de la fuente: 0.614 bits. Entropía máxima: 1 bit.

Observamos que la entropía de la fuente es menor que la máxima; las transmisiones *unicast* son significativamente más probables que las *broadcast*. Esto nos indica que los protocolos de control tienen un bajo impacto en la *performance* de la red.

2.5.2. Estructura de la red en base a los paquetes ARP

En la figura 7 se puede ver el grafo de la red subyacente de mensajes ARP.

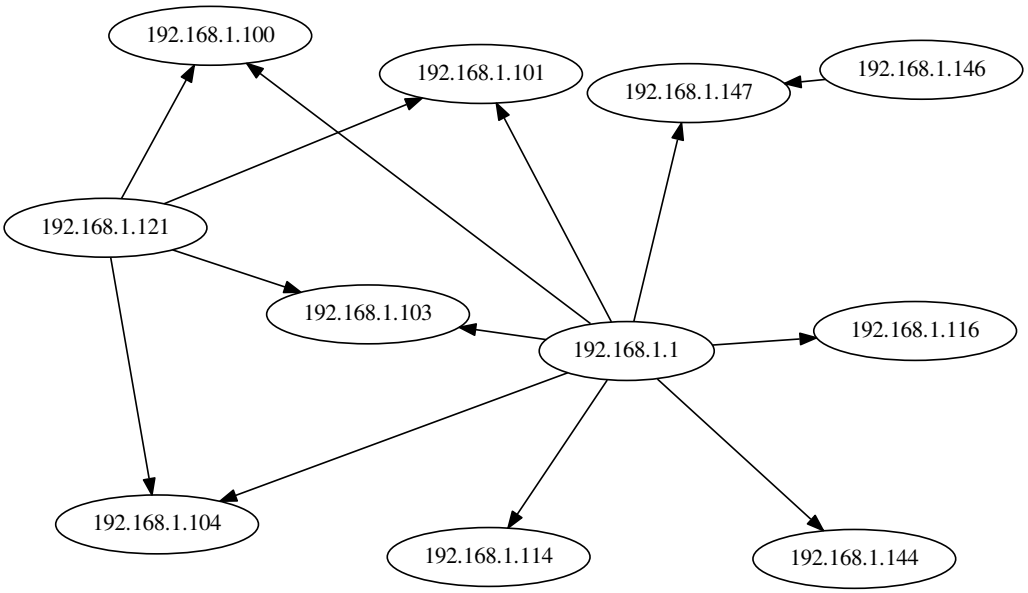

Figura 7: Grafo de la red subyacente de mensajes ARP en el experimento 3.

Todas las direcciones IP de la red son de la forma 192.168.1.X. Éstas son direcciones IP privadas[19].

En el grafo se destaca claramente un nodo, el 192.168.1.1. El vértice 192.169.1.121 también parece destacarse. Una comparación entre las MAC addresses que acompañan a estas dos direcciones en los paquetes ARP y los diversos dispositivos de la red nos indicó que la 192.168.1.1 corresponde al router de la red, mientras que la 192.168.1.121, a la interfaz de una de las computadoras.

2.5.3. Fuente S_1

En la figura 8 podemos ver la información de los mensajes de la fuente S_1, sin paquetes repetidos y tomando sólo el *Sender's Protocol Address* de los paquetes.

La entropía de la fuente es de 1.24 bits, siendo la máxima 32 bits.

La fuente clasifica exactamente de la forma deseada: el único nodo distinguido es el correspondiente al router.

En la figura 9 podemos ver la información de los mensajes de la fuente S_1, sin paquetes repetidos y tomando tanto el campo *Sender's Protocol Address* como el *Target Protocol Address* de los paquetes.

La entropía de la fuente es de 3.09 bits, siendo la máxima 32 bits.

Al igual que en los otros experimentos, la entropía es mayor al considerar esta fuente.

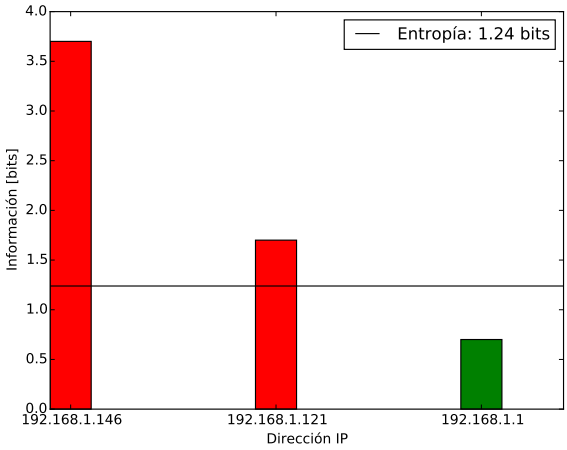

Figura 8: Información de los nodos de la fuente S_1 en el experimento 3, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP en el campo *Sender's Protocol Address* de un paquete ARP *who-has*.

[19] Todo el rango 192.168.0.0-192.168.255.255 es privado.

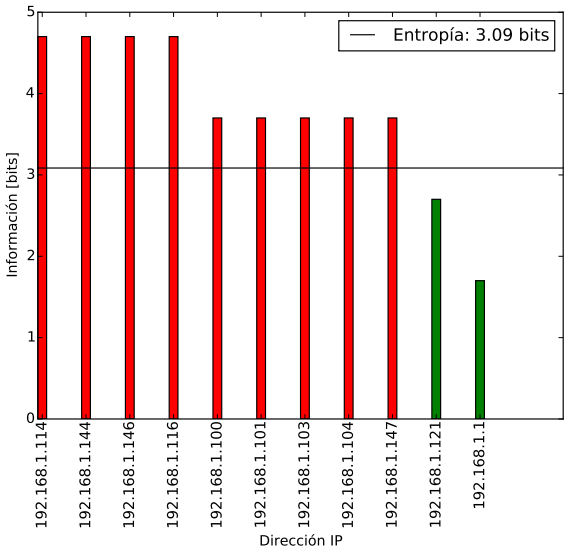

Figura 9: Información de los nodos de la fuente S_1 en el experimento 3, sin tomar paquetes repetidos y considerando como mensaje la ocurrencia de una IP tanto en el campo *Sender's Protocol Address* como en el *Target Protocol Address* de un paquete ARP *who-has*.

Se presenta un falso negativo, el previamente mencionado nodo 192.168.1.121. Este resultado es interesante ya que éste actúa exclusivamente como emisor, por lo que la cantidad de veces que aparece en ambas fuentes es igual; la mayor entropía de esta fuente lleva a que sea clasificado como vértice distinguido.

3. Conclusiones

La entropía de la fuente S no resultó la mejor herramienta para juzgar el efecto de los protocolos de control sobre la red, sino las probabilidades de las dos formas de transmisión[20].

Vimos que dicho efecto no depende exclusivamente del tamaño de la red; el grado de interconexión de los nodos juega un papel fundamental. Esto se evidencia en el hecho de que la transmisión *broadcast* era significativamente más probable en la red de trabajo que en la de los laboratorios del DC, a pesar de que la primera era más chica que la segunda.

Observamos que la entropía incrementa al aumentar el tamaño de la red, pero cabe destacar que el grado de interconexión de los nodos nuevamente juega un papel importante. Esto es de esperar, ya que la fuente de máxima entropía es una en el que todo mensaje es equiprobable; su grafo de ARP sería un grafo completo[21].

Los resultados al contemplar mensajes repetidos para la fuente S_1 fueron extremadamente pobres, con un muy alto grado de resultados erróneos, al punto que los nodos distinguidos no eran representativos de ninguna propiedad de la red. Por otro lado, tanto al incluir al destino de los paquetes ARP en la fuente como al no incluirlo, se obtuvieron resultados muy cercanos a lo buscado.

Ya que la mayoría de los nodos destacados principalmente actúan como emisores de los mensajes ARP, utilizar sólo el origen provee una buena aproximación. Utilizar ambos distingue ciertos nodos destacados que no lo fueron por la otra fuente, sin embargo la mayor entropía que presenta lleva a que nodos no destacados sean clasificados positivamente.

Es decir, tomar sólo el origen causa más falsos negativos, mientras que tomar ambos causa más falsos negativos. La diferencia de errores de clasificación entre ambas fuentes probó ser pequeña en nuestros experimentos, por lo que no recomendamos una por sobre la otra; concluimos que la decisión de fuente debería realizarse en base a cuál de estos dos errores se considere más leve.

Los resultados fueron satisfactorios para redes con una entropía en relación al tamaño relativamente baja[22], mientras que al aumentar esta magnitud se tornaron menos aceptables.

Las topologías de las redes modeladas a partir de los paquetes ARP se exhibieron sumamente diferentes. En particular, la red de los laboratorios del DC se mostró fragmentada, con las diversas componentes conexas presentando en la mayoría de los casos una estructura de estrella o similar; la red del trabajo se manifestó conexa con la excepción de dos nodos, con un grado de interconexión mucho más alto. Esta diferencia debe estar relacionada a la distinta función de las redes: la del DC busca brindar una conexión de internet a los diversos visitantes de los laboratorios, que buscarán acceder direcciones mayoritariamente fuera de la red; la del trabajo busca reunir además diversos dispositivos interdependientes.

[20] La entropía está dada por las probabilidades, pero esconde cuál forma de transmisión resulta más común; este dato es de suma importancia para el análisis planteado.

[21] Es decir, al incrementar la interconexión de los nodos, más se acerca a una fuente equiprobable de entropía máxima.

[22] Es decir, cuyo grafo de ARP es relativamente esparso, como en el caso de los experimentos 1 y 3.