Taller 1 - Wiretapping

Integrantes

LU	Nombre	Mail
72/18	Manuel Panichelli	panicmanu@gmail.com
76/16	Luciano Strika	lucianostrika44@gmail.com
692/12	Elías Cerdeira	eliascerdeira@gmail.com

Introducción

En este trabajo se elabora un programa que modela el tráfico de una red con dos fuentes de memoria nula distintas. Primero, analizaremos todos los paquetes que circulan por la red, que serán distinguidos por el tipo de protocolo y la dirección destino (*UNICAST* o *BROADCAST*). Finalmente, analizaremos únicamente los paquetes ARP distinguiéndolos por direccion de destino y fuente para enumerar todos los *hosts* de una red.

Métodos y condiciones de los experimentos

Para la ejecución de los experimentos se partió del código provisto por la cátedra, extendiendo el método `mostrar_fuente`, agregándole el cálculo de la información de cada símbolo y la entropía de la fuente.

La ejecución se realizó para 15000 tramas en tres redes domésticas distintas el 1/05/2021 aproximadamente a las 20:00 (GMT-3).

- Red 1 (Manuel): Consta de una computadora de escritorio conectada por UTP, dos celulares y una televisión.
- Red 2 (Luciano): Solo consta de una notebook, el router y un celular conectados por Wi-Fi.
- Red 3 (Elías): Consta de 11 dispositivos. 2 notebooks, 3 televisores, 2 decodificadores, 1 impresora y 3 celulares conectado vía WiFi.

Distinción de *hosts*

Para distinguir los *hosts* conectados a una red determinada, bastará con hallar todas las IPs que participan de la comunicación en el medio compartido, es decir, todas las que envían o reciben paquetes ARP.

Para esto, se eligió modelar una fuente de memoria nula S2, donde cada símbolo representa cada IP de fuente en un paquete ARP. Además, se eligió modelar otra fuente de memoria nula, donde los símbolos son las IP de destino. Se observaron las distribuciones de ambas y cuántos nodos aparecían en cada una.

Resultados de los experimentos

Resultado de análisis de protocolos para cada red.

1. Red de Manuel

   Entropía: 0.0403

Tipo de mensaje	Protocolo	Probabilidad	Información
UNICAST	2048 (IPv4)	0.99640	0.00520
BROADCAST	2054 (ARP)	0.00220	8.82828
BROADCAST	2048 (IPv4)	0.00073	10.41324
UNICAST	35130 (IEEE 1905.1a)	0.00027	11.87267
UNICAST	35020 (LLDP)	0.00027	11.87267
UNICAST	2054 (ARP)	0.00013	12.87267

2. Red de Luciano

   Entropía: 0.0093

Tipo de mensaje	Protocolo	Probabilidad	Información
UNICAST	2048 (IP)	0.99920	0.00115
UNICAST	2054 (ARP)	0.00080	10.28771

3. Red de Elías

   Entropía: 2.111

Tipo de mensaje	Protocolo	Probabilidad	Información
UNICAST	2048 (IPv4)	0.96774	0.047
UNICAST	34525 (IPv6)	0.01453	14.1
BROADCAST	33024 (IEEE 802.1Q VLAN)	0.00640	16.83
BROADCAST	2054 (ARP)	0.00507	17.6
UNICAST	33024 (IEEE 802.1Q VLAN)	0.00380	18.56
UNICAST	2054 (ARP)	0.00120	22.4
BROADCAST	2048 (IPv4)	0.00113	22.6
BROADCAST	34999 (OUI EE)	0.00013	29.73

Viendo estos datos, podemos responder algunas de las incógnitas planteadas.

- **¿Considera que las muestras obtenidas analizadas son representativas del comportamiento general de la red?**

  Se considera que la muestra tomada es representativa de las tres redes estudiadas, ya que predominan los paquetes IP y ARP.

- **¿Hay alguna relación entre la entropía de las redes y alguna característica de las mismas (ej.: tamaño, tecnología, etc)?**

  A mayor tamaño de la red, es decir, mayor cantidad de dispositivos simultáneamente conectados, y mayor diversidad tecnológica (e.g. usar IPv6) mayor será la entropía.

- **¿En alguna red la entropía de la fuente alcanza la entropía máxima teórica?**

  En una fuente de memoria nula la entropía máxima se alcanza cuando todos los símbolos que emite son equiprobables. Esto no sucede en la redes estudiadas, dado que siempre existe un símbolo que predomina.

- **¿Considera significativa la cantidad de tráfico *broadcast* sobre el tráfico total?**

  La cantidad de tráfico *broadcast* es despreciable en comparación al *unicast*.

- **¿Cuál es la función de cada uno de los protocolos encontrados? ¿Cuáles son protocolos de control y cuáles transportan datos de usuario? ¿Ha encontrado protocolos no esperados? ¿Puede describirlos?**

  Los protocolos esperados encontrados fueron IPv4, IPv6 (Transportan datos de usuario) y ARP (Protocolo de control). Luego se encontraron estos protocolos que no se esperaba observar que fueron buscados en IANA IEEE 802 Numbers y Wireshark.

  - `35020` (LLDP) Link Layer Discovery Protocol, usado por dispositivos para darse a conocer en la LAN.
  - `35130` (IEEE 1905.1a) Protocolo usado para redes domésticas (con soporte para wireless).
  - `33024` (IEEE 802.1Q VLAN) Estándar de redes para VLAN en Ethernet.
  - `34999` (OUI EE)
  - `35020` (LLDP): IEEE Std 802.1AB - Link Layer Discovery Protocol

Resultados de experimentos de distinción de *hosts*

A continuación se presenta el comportamiento observado al sniffear los paquetes ARP en las tres redes descritas al utilizar las IP cómo símbolos de la fuente.

1. Red de Manuel

- Fuente (Entropía: 0.3625)

IP	Probabilidad	Información
192.168.0.4	0.95000	0.07400
192.168.0.8	0.02000	5.64386
192.168.0.6	0.02000	5.64386
192.168.0.1	0.01000	6.64386

Destino (Entropía: 0.3625)

IP	Probabilidad	Información
192.168.0.1	0.95000	0.07400
192.168.0.10	0.02000	5.64386
192.168.0.6	0.02000	5.64386
192.168.0.4	0.01000	6.64386

2. Red de Luciano

- Fuente (Entropía: 1.0)

IP	Probabilidad	Información
192.168.0.1	0.50000	1.00000
192.168.0.181	0.50000	1.00000

- Destino (Entropía: 1.0)

IP	Probabilidad	Información
192.168.0.1	0.50000	1.00000
192.168.0.181	0.50000	1.00000

3. Red de Elías

- Fuente (Entropía: 0.9815)

IP	Probabilidad	Información
192.168.1.54	0.63000	0.46204
192.168.1.1	0.26000	1.34707
192.168.1.51	0.08000	2.52573
192.168.1.201	0.01000	4.60517
192.168.1.200	0.01000	4.60517
192.168.1.39	0.01000	4.60517

- Destino (Entropía: 1.4330)

IP	Probabilidad	Información
192.168.1.1	0.66000	0.41552
192.168.1.51	0.10000	2.30259
192.168.1.54	0.03000	3.50656
192.168.1.200	0.03000	3.50656
192.168.1.41	0.02000	3.91202

IP	Probabilidad	Información
192.168.1.58	0.02000	3.91202
192.168.1.65	0.02000	3.91202
192.168.1.33	0.02000	3.91202
192.168.1.52	0.02000	3.91202
192.168.1.64	0.02000	3.91202
192.168.1.201	0.02000	3.91202
192.168.1.39	0.02000	3.91202
192.168.1.63	0.01000	4.60517
192.168.1.36	0.01000	4.60517

- **¿La entropía de la fuente es máxima? ¿Qué sugiere esto acerca de la red?**

  Como se explicó anteriormente, la entropía máxima se da cuando la distribución de los símbolos es uniforme. Esto no se observa en la redes 1 y 3, lo cual sugiere que no todos los nodos están participando de forma equitativa en la resolución de direcciones. Mientras tanto, sí se observa un nivel de entropía máxima para la red 2, lo que probablemente se pueda adjudicar a la presencia de pocos nodos.

- **¿Se pueden distinguir nodos? ¿Se les puede adjudicar alguna función específica?**

  Sí, se pueden distinguir los nodos de las redes a través de las IPs.

  Para las redes 1 y 3, creemos que se le puede adjudicar al nodo con mayor proporción de tráfico entrante la función de *router*. Esto es porque para que los paquetes salgan de la red deben pasar por el *router*, por lo tanto los dispositivos necesitan conocer su dirección MAC. En cambio, para la red 2, la distribución es uniforme.

- **¿Hay evidencia parcial que sugiera que algún nodo funciona de forma anómala y/o no esperada?**

  Sí, para las redes 1 y 3 se observa un nodo que presenta la mayor parte del tráfico saliente, lo cual indicaría que es aquel que más frecuentemente requiere volver a mapear las direcciones. Creemos que esto puede deberse a que presenta una memoria caché de tamaño pequeño o un TTL corto en comparación a los demás nodos.

- **¿Existe una correspondencia entre lo que se conoce de la red y los nodos distinguidos detectados por la herramienta?**

  Sí, se detectó casi exactamente lo esperado, dado el conocimiento de la topología de las redes.

- **¿Ha encontrado paquetes ARP no esperados? ¿Se puede determinar para que sirven?**

  No, los paquetes ARP encontrados tienen operaciones 1 (*request*) y 2 (*response*). En la red 2, hay igual cantidad de ambos tipos de paquetes. Sin embargo, en las redes 1 y 3 esto se vuelve muy asimétrico, pues existen mas paquetes de *request*. No se logró dilucidar a qué se debe este fenómeno.

Conclusiones

El trabajo no presentó dificultades significativas, aunque generó dudas respecto de la falta de paquetes ARP en broadcast para la red número 2, que aún no se pudieron resolver.

Además, como era esperado, se observa que la mayor parte del tráfico termina siendo IPv4 *unicast*.

Al sniffear el tráfico de paquetes de las redes a través de Wireshark, nos resultó llamativa la cantidad elevada de paquetes UDP con respecto a la de TCP. Creemos que esto puede deberse a la transferencia de datos multimedia generada al realizar *streaming* de audio y video a través de Discord.

Finalmente, nos resultó sumamente interesante que mediante un análisis tan simple del tráfico de paquetes ARP se pueda conocer la topología de una red. Probablemente valdría la pena realizarlo en una red pública de tamaño mayor para ver cómo se comporta.