Trabajo Práctico №1: Wiretapping

Alvaro Jose Fernando, Barbeito Nicolás, Brum Raúl, Nievas Yésica

***Resumen*—Our premise is ...**

I. Introducción

A. Paquetes ARP

El protocolo ARP (Address Resolution Protocol) permite mapear direcciones de nivel de red a direcciones físicas. La idea de este protocolo se basa en el envío de paquetes que pueden ser de preguntas o respuestas. El emisor del paquete que pregunta por una dirección, envía un mensaje broadcast sobre la red local, siendo respondido por un mensaje unicast por aquel al que pertenece la dirección consultada. Mediante el envío de estos paquetes ARP se construyen las tablas que mapean direcciones de red con direcciones físicas.

B. Entropía de una fuente

Para poder definir la entropía de una fuente, necesitamos la definición de información que aportan los símbolos emitidos por dicha fuente. Se define información de un símbolo s como

$$\mathrm{I}\ (s) = \log(1/\mathrm{P}(s))$$

siendo $\mathrm{P}(s)$ la probabilidad de ocurrencia de dicho símbolo. De esta manera, puede calcularse la información media suministrada por una fuente de información de memoria nula (los símbolos emitidos son estadísticamente independientes) como

$$\sum_S P(s_i)I(s_i)\forall s_i \in S$$

Esta cantidad media de información por símbolo de la fuente, recibe el nombre de *entropía* H(S) de la fuente de memoria nula.

$$H(S) = \sum_S P(s_i)log(1/P(s_i))$$

Debido a que la entropía de una fuente depende de la probabilidad de los diferentes símbolos que la componen, se puede demostrar que para una fuente de información de memoria nula con un alfabeto de q símbolos, el valor máximo de la entropía es precisamente log q, alcanzándose solamente si todos los símbolos son equiprobables.

C. Fuente S

Sea P la fuente de información generada a partir de todos los paquetes Ethernet que se transmiten en una determinada red entre los instantes de tiempo $[t_i, t_f]$:

$P_{ti,tf} = \{p_1 ... p_n\}$

siendo p_i el i-ésimo paquete transmitido en la red entre los instantes de tiempo $[t_i, t_f]$.

Alvaro Jose Fernando, LU: 89/10, email: fer1578@gmail.com
Barbeito Nicolás, LU: 147/10, email: nicolasbarbeiton@gmail.com
Brum Raúl, LU: 199/98, email: brumraul@gmail.com
Nievas Yésica, LU: 340/05, email: yesica.nievas@gmail.com

Los paquetes p_i pertenecientes a P encapsulan diferentes protocolos, que se pueden identificar a través del campo type del frame de capa 2 (p.type en Scapy). Por lo tanto, con el objetivo de distinguir los protocolos utilizados en una red, se define otra fuente de información S de la siguiente manera:

$S_{ti,tf} = \{s_1 ... s_n\}$

siendo $\mathrm{s}_i = \mathrm{p}_i$.type /$\mathrm{p}_i$ perteneciente a P entre los instantes de tiempo $[t_i, t_f]$.

D. Propuesta de una nueva fuente S_1

Para analizar la entropía de la red en base a los paquetes ARP observados realizamos una nueva tool en base a la anterior que nos permitiera obtener datos de los campos de dichos paquetes. Se propone como nueva fuente S1 el conjunto de símbolos conformado por las distintas direcciones IP destino:

$\mathrm{S1} = \{\mathrm{s1}_1 ... \mathrm{s1}_n\}$ siendo $\mathrm{s1}_i$ el valor del campo *pdst* correspondiente a la ip destino del paquete

Al igual que para la fuente S, realizamos el cálculo de la entropía como fue requerido, como la probabilidad e información de sus símbolos.

II. Métodos y condiciones de cada experimento

En esta sección describiremos brevemente las redes elegidas para realizar las escuchas mediante las herramientas indicadas en el enunciado del trabajo practico.

A. Home Lan

Esta medición fue realizada en la Lan de una casa por un intervalo de 2 horas. La misma cuenta con una computadora corriendo un sistema operativo Linux la cual es el router de la Lan y provee a las demas computadoras de acceso a internet ademas de otros servicios de red (proxy, dns, dhcpcd, etc). A la misma se encuentran conectadas mediante un switch 4 computadoras cableadas y 2 access point inalambricos. A estos últimos se encontraban conectados al momento de la medición una notebook y varios celulares. Ademas una de las computadoras cableadas corre una maquina virtual con ip propia independiente y otra de las computadoras cableadas posee otro isp para conectarse a internet por lo que no utiliza a la primera computadora como gateway.

B. red 2

C. red 3

D. red 4

III. Resultados y Análisis

En esta sección mostraremos y analizaremos los resultados obtenidos en las mediciones realizadas en las distintas redes tanto para la fuente S y S_1 descriptas en la sección I C y D. Para cada una de las redes veremos sus protocolos distinguidos, la proporción de paquetes ARP en el trafico de la red y los nodos (representados por ip) distinguidos.

A. Home Lan

A.1 Fuente S

Los resultados obtenidos para la fuente S fueron:

Protocolo	Informacion	Probabilidad
EAPOL	12.31	0.01 %
ARP	5.11	2.88 %
IPv6	4.22	5.33 %
IPv4	0.12	91.76 %

TABLE I
Home Lan - Protocolos

Como puede observarse en la figura 1 el protocolo mas utilizado es IPv4 en un 91.76 % mientras que IPv6 y ARP solo son utilizados en un 5.33 % y 2.88 % respectivamente. EAPOL (autenticación wireless) no tiene prácticamente incidencia. Observamos ademas que el protocolo ARP tiene solo una incidencia del 2.88 % en el trafico total de la fuente lo que hace que el overhead aportado por el mismo no sea significativo.

Figura 1. Home Lan - Probabilidades

Por otro lado la entropía de la fuente S fue de 0.4892 lo que hace que los símbolos emitidos por la fuente S sean muy previsibles. Esto podemos notarlo en la figura 2 donde observamos que el protocolo con mayor porcentaje de aparición, IPv4, sea el que menos información aporta a la fuente. La información aportada por este se encuentra por debajo de la entropía de S. Como contraste observamos en la figura 2 que el protocolo EAPOL es el que mas información aporta pero según lo observado en la figura 1 tiene una probabilidad muy baja lo que hace que no incida en la entropía.

Figura 2. Home Lan - Información

A.2 Fuente S_1

A continuación se adjuntan algunos los resultados obtenidos:

- **"Muestra 1"*(captura1.pcap)*:**
  - `Entropía Fuente S:`
  - `Entropía Fuente S1:`
- **"Muestra 2"*(captura2.pcap)*:**
  - `Entropía Fuente S:`
  - `Entropía Fuente S1:`
- **"Muestra 3"*(captura3.pcap)*:**
  - `Entropía Fuente S:`
  - `Entropía Fuente S1:`
- **"Muestra 4"*(captura4.pcap)*:**
  - `Entropía Fuente S:`
  - `Entropía Fuente S1:`

B. Protocolos y nodos distinguidos y proporción de paquetes ARP

IV. Gráficos y análisis

A continuación haremos un análisis de cada una de las redes de las cuales capturamos su tráfico. Para ello, haremos uso de distintos gráficos que nos ayudarán a visualizar mejor la información del tráfico en cada red.

A. Analizando tráfico de çaptura1"(captura1.pcap)

B. Analizando tráfico de çaptura2"(captura2.pcap)

C. Analizando tráfico de çaptura3"(captura3.pcap)

D. Analizando tráfico de çaptura4"(captura4.pcap)

V. Referencias