DEPARTAMENTO DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico 1

Wiretapping

3 de diciembre de 2015 — Teoría de las comunicaciones

Integrante	LU	Correo electrónico
Ladelfa, Hernán Nahuel	318/04	nahueladelfa@gmail.com
González, Sergio Martín	723/10	sergiogonza90@gmail.com

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 2610 - C1428EGA
Ciudad Autónoma de Buenos Aires - Rep. Argentina
Tel/Fax: (++54 +11) 4576-3300
http://www.exactas.uba.ar

Índice

1. Introducción

El objetivo de este trabajo, es poder realizar un estudio sobre distintas redes, para poder realizar un análisis de los distintos **protocolos** que corren en la misma, asi como la posible caracterización de algunas IPs, como **nodos distinguidos**

Para realizar esto, se implemento una herramienta capas de: poder capturar el trafico de una red, ya sea Wifi o Ethernet 802.3, y quedarse solo con la información de los distintos protocolos encontrados, y los distintos hosts (IP/MAC) que intervienen.

La herramienta también permite calcular la entropia, y la probabilidad de aparición de un nodo, en el trafico de la red en un período de tiempo determinado.

Con esta información, el objetivo es realizar un análisis exhaustivo de las redes estudiadas.

2. Resumen de la herramienta

Para la simulación de las fuentes de información anteriormente presentadas. Desarrollamos una aplicación utilizando Scapy, la cual escucha pasivamente la red, captura los paquetes y vuelva la información pertinente en una serie de archivos. Los mismos son los siguientes:

- Un archivo que indica: El origen y el destino (MAC) de la capa Ethernet, el "type", y las IPs origen y destino (solo cuando es pertinente)
- Un archivo con la probabilidad de ocurrencia de cada "type", y la entropía de los "type"
- Un archivo con la probabilidad de ocurrencia de cada "host" (ip), y la entropía de los host (solo teniendo en cuenta paquetes ARP)
- Una imagen con un grafo indicando los "request" (who-has) y "replys" (is-at) que se enviaron los "host" entre si (opcional).

2.1. Ejecución

La ejecución del a herramienta debe ser realizada en un entorno Linux, y es necesario tener instalado: *Python 3.0* (o superior), *Scapy*, y si ademas se quiere realizar grafos, es necesario tener instalado *graphviz*.

Para ejecutar la herramienta, se debe abrir una consola en la carpeta "src" adjunta a este informe y ejecutar el comando:

```
$ sudo ./WiretappingTool.py
```

Esto comenzara a capturar, y volcara la informacion en los archivos configurados como default. Adicionalmente, la aplicacion cuenta con ciertos paramentos para personalizar la ejecucion de la captura. Por ejemplo:

```
$ sudo ./WiretappingTool.py -f salida.out -t 60 --console --arp
```

Esto realiza lo sigiuente:

- (-t o --timeout) indica un tiempo en segundos para finalizar la captura, en esta caso 60 segundos.
- (--arp) Captura solo paquetes ARP.
- (--console) Ademas de volcar el resultado de la captura a un archivo, se muestran los paquetes en la consola, en tiempo real.
- (-f) Vuelca la captura en *salida.out* en lugar del archivo configurado por defecto (out/sniff.out).

Si ademas de la captura, se desea realizar un grafo con los nodos de la red, junto con los "request" y "replys" realizados por los mismos, se debe pasar el parametro "–graph". Ejemplo:

```
$ sudo ./WiretappingTool.py -f salida.out --console --arp --graph
```

Al terminar la ejecucion, la imagen con el grafico se abrirá al instante.

Para ver con detalle cada uno de los parametros que dispone la aplicacion, se puede utilizar:

```
$ sudo ./WiretappingTool.py -h
```

3. Primera Consigna: Caputando tráfico

3.1. Desarrollo

3.2. Fuentes de información

3.2.1. Protocolos

En primer lugar, la cátedra planteo una fuente de información que nos ayude a distinguir entre los distintos protocolos que se pueden hallar en una red. La fuente de información es la siguiente:

$$S_{t_i;t_f} = \{s_1 \cdots s_n\} \text{ siendo } s_i = p_i.type / p_i \in P \text{ entre los instantes de tiempo } [t_i; t_f].$$

y P esta definida de la siguiente manera:

$$P_{t_i;t_f} = \{p_1 \cdots p_n\} \text{ siendo } p_i \text{ el i-esimo paquete transmitido en la red entre los instantes de tiempo } [t_i, t_f]$$

Entonces, lo que se hace es, realizar una captura de la red, entre los instantes t_i y t_f, y de esto nos quedamos únicamente con los campos *type* del frame Ethernet. Así obtenemos una fuente de informacion donde cada simbolo es un protocolo utilizao en la red.

3.2.2. Nodos

Ahora, debemos idear una fuente de información, que nos permita poder distinguir los hosts involucrados en la captura de trafico, y a su ves, distinguir entre ellos, cuales son los mas concurridos y con mas apariciones en la red.

Para esto, tomemos la fuente de información P vista en el anterior punto, y tratemos de adaptarla para poder obtener una fuente de información que nos permita encontrar hosts en la red, solo basándonos en paquetes ARP. Tomemos el siguiente subconjunto de P:

$$\overline{P}_{t_i;t_f} = \{p_1 \cdots p_n\} \; \forall \; p_i \in \text{P} \; / \; p_i.type = ARP \text{ entre los instantes de tiempo } [t_i; t_f].$$

Con esta fuente de información, lo que hacemos es quedarnos únicamente con los paquetes ARP de toda la captura de trafico. En base a esto proponemos la siguiente fuente de información.

$S_1 = R_{t_i;t_f} = \{r_{ai} \mid r_{ai} = \overline{p}_i[ARP].ip\_origen\} \bigcup \{r_{bi} \mid r_{bi} = \overline{p}_i[ARP].ip\_destino\}$ entre los instantes de tiempo $[t_i; t_f]$.

Esto quiere decir, nos quedamos con las IPs origen y destino de la capa ARP del paquete (del cual sabemos que tiene una, porque los símbolos son tomados de $\overline{P}$). De esta forma, obtenemos una fuente de información donde cada simbolo es un hosts de la red.

Teniendo esto, podemos calcular, tanto la entropía de ambas fuentes, así como la probabilidad y cantidad de información de cada símbolo, tanto para S_1 como para $S_{t_i;t_f}$. La cantidad de información de cada símbolo s, se calcula de la siguiente forma:

$$I(s) = -log_2(P(s))$$

en donde $P(s)$ es la probabilidad de ocurrencia de s, en todo el espacio muestral S, en este caso, todos los símbolos emitidos por S_1 y $S_{t_i;t_f}$. Luego, la entropía de la fuente es la que se obtiene mediante la siguiente formula:

$$H(S) = \sum_{s \in S} P(s) * I(s)$$

4. Segunda Consigna: Gráficos y Análisis

5. Conclusiones