TP1: Wiretapping

Teoría de las Comunicaciones
Departamento de Computación
FCEN - UBA

30.03.2016

1. Introducción

El objetivo de este trabajo es utilizar técnicas provistas por la teoría de la información para distinguir diversos aspectos de la red de manera analítica. Además, sugerimos el uso de dos herramientas modernas de manipulación y análisis de paquetes frecuentemente usadas en el dominio de las redes de computadoras: Wireshark [?] y Scapy [?].

2. Normativa

- Fecha de entrega: 20-04-2016.
- El informe deberá haber sido enviado por correo para esa fecha con el siguiente formato:
  to: tdc-doc at dc uba ar
  subject: debe tener el prefijo [tdc-wiretapping] y contener en numero de grupo
  body: nombres de los integrantes y las respectivas direcciones de correo electrónico
  attachments: el informe en formato pdf + el código fuente en formato zip. Junto con los fuentes debe haber un archivo de texto con indicaciones para su uso o un makefile en su defecto.
- No esperar confirmación. Todos los mails llegan a la lista a menos que reciban una respuesta indicando explícitamente que el mail fue rechazado. Los avisos por exceso de tamaño no son rechazos.

3. Enunciado

3.1. Introducción

Sean $p_1..p_i$ los paquetes que se transmiten en un enlace. Podemos conocer los protocolos que encapsulan los paquetes con el campo type del frame de capa de enlace (p_i.type en Scapy).

Se define el conjunto de símbolos $S = \{s_1 \cdots s_n\}$ siendo $s_i = p_i.type \ / \ p_i \in P$ entre los instantes de tiempo $[t_i, t_f]$. Podemos pensar la fuente S que tiene como símbolos a los protocolos utilizados en la red en ese intervalo.

Por último, en el marco de una fuente de información, podemos pensar a un **símbolo como distinguido** cuando sobresale del resto en términos de la información que provee.

3.2. Primera consigna: capturando tráfico

1. Implementar una herramienta que escuche pasivamente los paquetes Ethernet de la red.
2. Adaptar la herramienta del punto anterior para calcular la entropıa de la fuente S en la red local.
3. Proponga una fuente de información S1 con el objetivo de distinguir, en lugar de los protocolos como hace S, los nodos (hosts) de la red. La distinción de S1 debe estar basada únicamente en paquetes que utilicen el protocolo ARP y el criterio para la diferenciación lo deberá establecer el grupo.

3.3. Segunda consigna: gráficos y análisis

Utilizando estas herramientas, realizar experimentos para analizar:

- Los protocolos distinguidos.
- La incidencia de paquetes ARP en la red.
- Los nodos distinguidos.

Se deben realizar tantos experimentos como cantidad de miembros tenga el grupo. Además, las capturas deben ser lo más extensas posibles (tf - ti >10 minutos). En la medida de lo posible, intentar capturar en al menos una red que no sea controlada (trabajo, shopping, etc).

Los análisis deben estar basados en conceptos formales de la teoría de la información. O sea, se debe analizar qué símbolos son significativos en cada red, viendo la diferencia entre su información y la entropía de la fuente.

Por último, el informe debe seguir la siguiente estructura: somera introducción, métodos y condiciones de cada experimento, resultados y conclusión. La presentación de los resultados debe efectuarse mediante gráficos y su correspondiente análisis. Sugerimos, entre otros, histogramas (de IPs y protocolos) con cortes en los valores de entropía. Se valorará especialmente en esta consigna la creatividad y el análisis propuesto. Recomendamos entonces pensar cómo resultará más efectivo presentar la información recopilada.

Referencias

[1] RFC 826 (ARP) `http://tools.ietf.org/html/rfc826`

[2] Wireshark (página web oficial) `http://www.wireshark.org`

[3] Scapy (página web oficial) `http://www.secdev.org/projects/scapy/`

[4] OUI (IEEE) `http://standards.ieee.org/develop/regauth/oui/oui.txt`