1 Objetivos

- Aplicar los conocimientos sobre redes de computadoras en el análisis estadístico de tráfico de red, para ganar conocimiento del comportamiento de la red y detectar anomalías.
- Aplicar los conocimientos sobre detección de intrusos, reducción de dimensiones, y otros temas vistos en clase para la construcción de un IDS.

2 Preámbulo

Las herramientas sniffer (pasivas) como Wireshark permiten la captura de tráfico de red para su posterior monitoreo y análisis. Con esta información se pueden descubrir aspectos de la red en lo referente a eficiencia y optimización (cuellos de botella), pero también se pueden detectar anomalías que sugieren un posible ataque.

3 Desarrollo

Parte 1 - Análisis de paquetes

Para este ejercicio se utilizará el archivo analisis_paquetes.pcap que se encuentra en CANVAS.

Análisis estadístico

1. Capture 10 paquetes con la herramienta scapy, imprima el tipo de datos, la longitud y el contenido del pcap
2. Añada al pcap de 10 paquetes el archivo analisis_paquetes.pcap
3. Convierta el pcap a un DataFrame
4. Muestre los valores de las columnas: Src Addres, Dst Address, Src Port y Dst Port
5. Estadísticas
    a. Muestre cual es la IP origen más frecuente
    b. Muestre cual es la IP destino más frecuente
    c. ¿A qué IPs se comunica la IP del inciso a?
    d. ¿A qué puertos destino se comunica la IP del inciso a?
    e. ¿A qué puertos origen se comunica la IP del inciso b?
    f. Indique el propósito de los puertos que más aparece en los incisos d y e
6. Gráficas
    a. Muestre una gráfica 2D, en el eje Y las IPs origen, y en el eje X la suma de los payloads enviados de dichas direcciones.
    b. Muestre una gráfica 2D, en el eje Y las IPs destino, y en el eje X la suma de los payloads recibidos en dichas direcciones.
    c. Muestre una gráfica 2D, en el eje Y los puertos origen, y en el eje X la suma de los payloads enviados de dichos puertos.

   d. Muestre una gráfica 2D, en el eje Y los puertos destino, y en el eje X la suma de los payloads recibidos en dichos puertos.
7. Investigación del payload
   a. Cree un nuevo DF que incluya únicamente las conexiones con la dirección IP origen más frecuente
   b. Obtenga un nuevo DF con las columnas Src Address, Dst Address y agrúpelas por payload
   c. Obtenga la IP que más ha intercambiado bytes con la IP más frecuente. Esta IP es sospechosa por la cantidad de bytes intercambiados, entre todas las direcciones.
   d. Cree un nuevo DF con la conversación entre la IP más frecuente y la IP sospechosa.
   e. Obtenga los payloads del DF del inciso 6, y añada cada uno en un array.
   f. Muestre el contenido del array.
   g. Examine los primeros bytes del contenido, ¿encuentra algún dato que no haga sentido que se envié a través del puerto explicado en el inciso f?

Parte 2 – Construcción de un IDS

- Descargue los archivos trafico_train.csv y trafico_test.csv proporcionados en CANVAS.
- En el archivo de entrenamiento se encuentra la columna Class que indica si el flujo de red es normal o una anomalía.
- Cree dos modelos para un IDS: un modelo que utilice PCA luego del pre procesamiento de datos, y un modelo sin reducción de dimensionalidad. Compare y explique las métricas (accuracy, precision y recall) de ambos modelos para los datos de entrenamiento y pruebas. Ambos modelos deben usar el mismo clasificador (no se puede utilizar decision trees o random forest).