Universidad de Granada

decsai.ugr.es

Inteligencia Artificial en Telecomunicaciones

Máster en Ingeniería de Telecomunicaciones

Práctica 3: Aprendizaje Automático con WEKA

DECSAI

Departamento de Ciencias de la Computación e Inteligencia Artificial

Detección de intrusiones de red

- En esta práctica se ha utilizado la herramienta de aprendizaje automático WEKA para diseñar un detector de intrusiones de red.
- Se puede definir intrusión como cualquier conjunto de acciones que tratan de comprometer la integridad, confidencialidad o disponibilidad de un recurso.
- La detección de intrusos es la capacidad de detectar ataques en una red, incluyendo dispositivos y computadores.

Origen de los Datos

- Los datos que se van a utilizar en este estudio son una pequeña selección del conjunto de datos del concurso KDD 1999, en donde se usó una versión reducida de la amplia variedad de intrusiones militares simuladas en un entorno de red, proporcionadas por DARPA Intrusion Detection Program Evaluation en 1998, que tenían como objetivo evaluar el estudio y la investigación en la detección de intrusiones.
- Los Laboratorios Lincoln crearon un entorno para adquirir un volcado de datos TCP durante nueve semanas, en una red de área local (LAN) que simulaba la típica red de las Fuerzas Aéreas de EE.UU salpicada con múltiples ataques.
- El conjunto bruto de datos de entrenamiento, obtenidos durante las primeras 7 semanas, ocupaban cerca de cuatro gigabytes, lo que equivale aproximadamente a cinco millones de registros de conexión.

Tipos de Ataques

- Los ataques se dividen en cuatro categorías principales:
  - DoS (denial-of-service): denegación de servicio. Hace que un servicio o recurso sea inaccesible a los usuarios legítimos. Normalmente provoca la pérdida de la conectividad de la red por el consumo del ancho de banda de la red de la víctima o sobrecargando los recursos computacionales de su sistema [9].
  - R2L: acceso no autorizado desde una máquina remota;
  - U2R: acceso no autorizado a privilegios de superusuario.
  - Probing: vigilancia y otros tipos de sondeo de redes.
- Algunos expertos creen que la mayoría de los ataques son variantes de ataques conocidos y la "firma" de estos pueden ser suficientes para capturar las nuevas variantes.
- Las bases de datos contienen un total de 24 tipos de ataque.

Tipos de Ataques

Ataque	Descripción	Tipo
back	Ataque contra el servidor web Apache cuando un cliente pide una URL que contiene muchas barras.	DoS
land	Envío de TCP/SYN falso con la dirección de la víctima como origen y destino, causando que se responda a sí mismo continuamente.	DoS
neptune	Inundación por envíos de TCP/SYN en uno o más puertos.	DoS
pod	Ping de la muerte: manda muchos paquetes ICMP muy pesados.	DoS
smurf	El atacante envía un ping, que parece proceder de la víctima, en broadcast a una tercera parte de la red, donde todos los host responderán a la víctima.	DoS
teardrop	Usa el algoritmo de fragmentación de paquetes IP para enviar paquetes corruptos a la víctima.	DoS
ftp_write	Usuario FTP remoto crea un archivo .rhost y obtiene un login local.	R2L
guess_passwd	Trata de adivinar la contraseña con telnet para la cuenta de visitante	R2L
imap	Desbordamiento remoto del búfer utilizando el puerto imap.	R2L
multihop	Escenario de varios días donde el atacante primero accede a una máquina que luego usa como trampolín para atacar a otras máquinas.	R2L
phf	Script CGI que permite ejecutar comandos en una máquina con un servidor web mal configurado.	R2L
spy	Analizador de protocolos LAN por la interfaz de red.	R2L
warezclient	Los usuarios descargan software ilegal publicado a través de FTP anónimo por el warezmaster.	R2L
warezmaster	Subida FTP anónima de Warez (copias ilegales de software).	R2L
buffer_overflow	Desbordamiento de la pila del búfer.	UR2
loadmodule	Ataque furtivo que reinicia la IFS para un usuario normal y crea un shell de root.	UR2
perl	Establece el id de usuario como root en un script de perl y crea un shell de root.	UR2
rootkit	Escenario de varios días donde un usuario instala componentes de un rootkit.	UR2
ipsweep	Sondeo con barrido de puertos o mandando pings a múltiples direcciones de host.	Probing
nmap	Escaneo de redes mediante la herramienta nmap.	Probing
portsweep	Barrido de puertos para determinar qué servicios se apoyan en un único host.	Probing
satan	Herramienta de sondeo de redes que busca debilidades conocidas.	Probing

Atributos

Tabla 3.3-1. Atributos básicos de las conexiones TCP.

Atributo	Descripción	Tipo
duration	Longitud (número de segundos) de la conexión.	Continuo
protocol_type	Tipo de protocolo (tcp...)	Discreto
service	Tipo de servicio de destino (HTTP, Telnet, SMTP…)	Discreto
src_bytes	Número de bytes de datos de fuente a destino	Continuo
dst_bytes	Número de bytes de datos de destino a la fuente.	Continuo
flag	Estado de la conexión (SF, S1, REJ…)	Discreto
land	1 si la conexión corresponde mismo host/puerto; 0 de otro modo.	Discreto
wrong_fragment	Número de fragmentos erróneos.	Continuo
urgent	Número de paquetes urgentes.	Continuo

Tabla 3.3-2. Atributos especiales.

Atributo	Descripción	Tipo
hot	Número de indicadores “hot”.	Continuo
num_failed_logins	Número de intentos de acceso fallidos.	Continuo
logged_in	1 si acceso exitoso; 0 de otro modo.	Discreto
num_compromised	Número de condiciones “sospechosas”.	Continuo
root_shell	1 si se obtiene superusuario para acceso a root; 0 de otro modo.	Discreto
su_attempted	1 si se intenta el comando “su root”; 0 de otro modo.	Discreto
num_root	Número de accesos a root.	Continuo
num_file_creations	Número de operaciones de creación de ficheros.	Continuo
num_shells	Número de Shell prompts.	Continuo
num_access_files	Número de operaciones de control de acceso a ficheros.	Continuo
num_outbound_cmds	Número de comandos de salida en una sesion ftp.	Continuo
is_hot_login	1 si el login pertenece a la lista “hot”; 0 de otro modo.	Discreto
is_guest_login	1 si el acceso es un “guest” 'login; 0 de otro modo.	Discreto

Atributos

Tabla 3.3-3. Atributos con ventana de dos segundos.

Atributo	Descripción	Tipo
count	Número de conexiones a la misma máquina que la conexión actual en los dos últimos segundos	Continuo
Los siguientes atributos se refieren a las conexiones de mismo host.		
serror_rate	Porcentaje de conexiones que tienen errores “SYN”.	Continuo
rerror_rate	Porcentaje de conexiones que tienen errores “REJ”.	Continuo
same_srv_rate	Porcentaje de conexiones con el mismo servicio.	Continuo
diff_srv_rate	Porcentaje de conexiones con diferentes servicios.	Continuo
srv_count	Número de conexiones al mismo servicio que la conexión actual en los dos últimos segundos	Continuo
Los siguientes atributos se refieren a las conexiones de mismo servicio.		
srv_serror_rate	Porcentaje de conexiones que tienen errores “SYN”.	Continuo
srv_rerror_rate	Porcentaje de conexiones que tienen errores “REJ”.	Continuo
srv_diff_host_rate	Porcentaje de conexiones a diferentes hosts.	Continuo

Datos

- Se proporcionan los siguientes conjuntos de datos:
  - KDDCup99.arff: Subconjunto del 10% de los datos (75 MB descomprimido).
  - KDDCup99_full.arff: Datos originales (430 MB descomprimido). No aconsejo su uso.
- Se debe tener en cuenta que la herramienta WEKA tiene que manejar grandes conjuntos de datos y que al usar la máquina virtual de java se dispone de una memoria limitada (usar java -jar-Xmx2048m weka.jar)

Ejercicio

- Probar tres clasificadores distintos y ver cual obtiene mejor resultados usando una validación cruzada de 10-hojas.
- Se aconseja usar sólo el conjunto de datos con el 10% de los datos.