Desarrollando sistemas de detección de intrusiones

Del laboratorio a su implantación

Facultad de Informática – Univ. Complutense de Madrid

Abril 2019

Jesús Esteban Díaz-Verdejo

Departamento de Teoría de la Señal, Telemática y Comunicaciones

E.T.S. Ingenierías Informática y Telecomunicación – Universidad de Granada

C/ Periodista Daniel Saucedo Aranda, s/n - 18071 – Granada (Spain)

Phone: +34-958 242304 - Email: jedv@ugr.es

© JEDV, 200575724

Esquema

Introducción
- Despliegue de la seguridad
- Desarrollo/investigación en IDS

Sistemas de detección de intrusiones
- Ataques
- Indicadores de compromiso
- Elementos
- Tipos de IDS

Desarrollo de un IDS
- Medidas de rendimiento
- Elementos de diseño

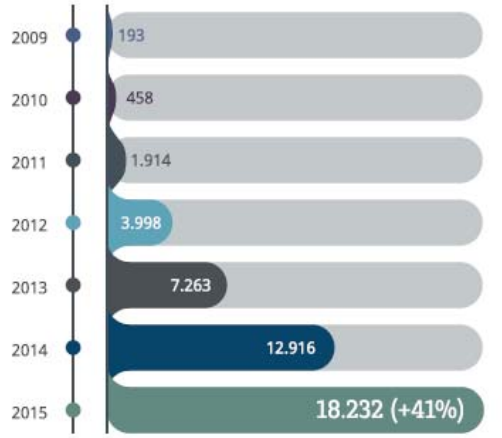
Deficiencias típicas
- Datasets
- Ejemplos: DARPA y KDD

Desafíos y futuro
- Aplicabilidad real
- Modelado de incidentes

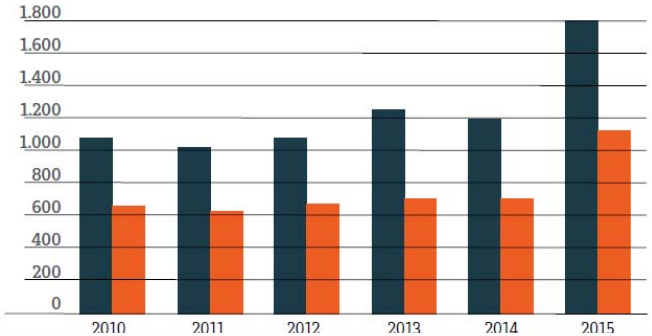
© 2019 - Jesús E. Díaz Verdejo

1. Introducción

- Importancia incuestionable de la seguridad en Internet
  - Aumento del número de incidentes de seguridad
    - Estadísticas de incidentes (Fuente: CCN-CERT):



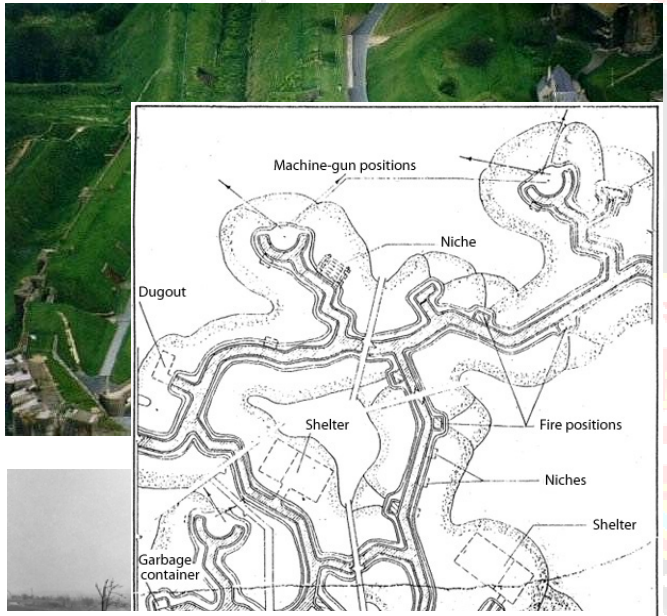
Evolución de los Incidentes gestionados por el CCN-CERT

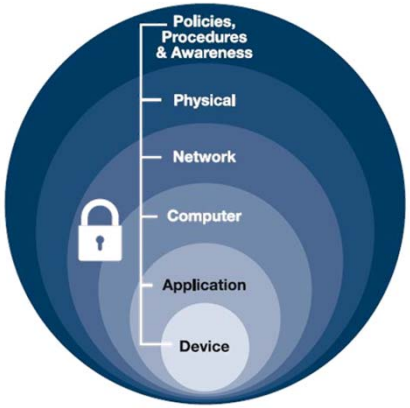


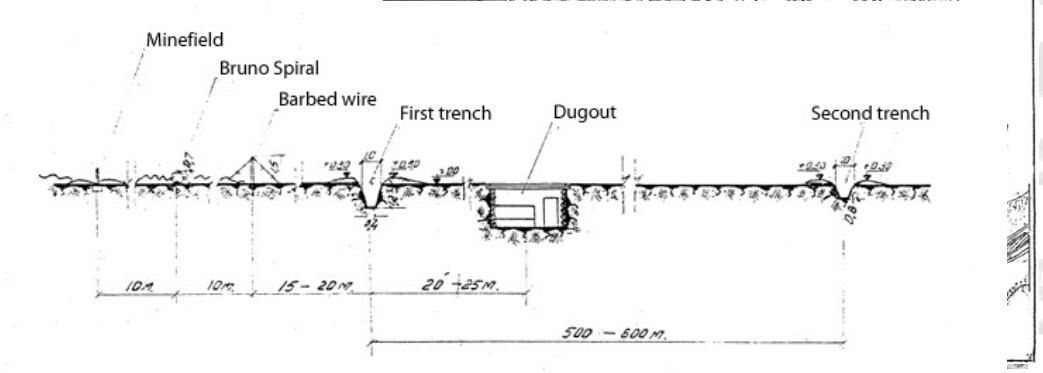
© 2019 - Jesús E. Díaz Verdejo

Despliegue de la seguridad

- Defensa en profundidad
  - Combinación de múltiple controles y mecanismos de mitigación independientes
    - Estrategia para retardar al atacante mientras es neutralizado
  - Despliegue de los mecanismos de defensa en **capas**

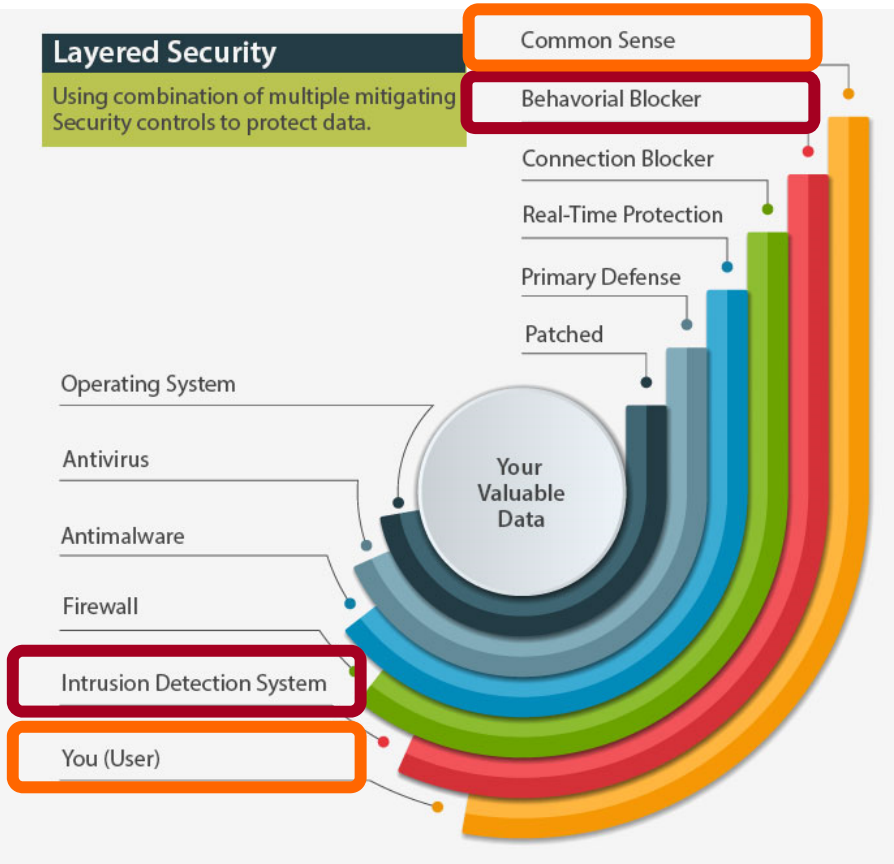




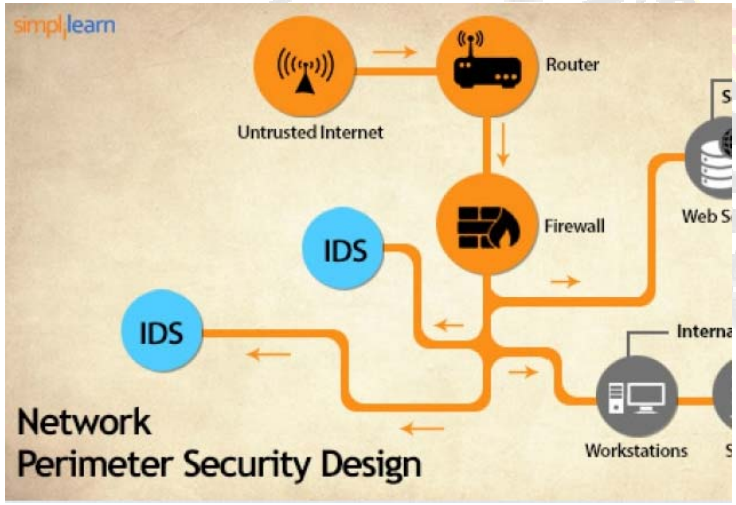


© 2019 - Jesús E. Díaz Verdejo

Elementos de defensa en profundidad



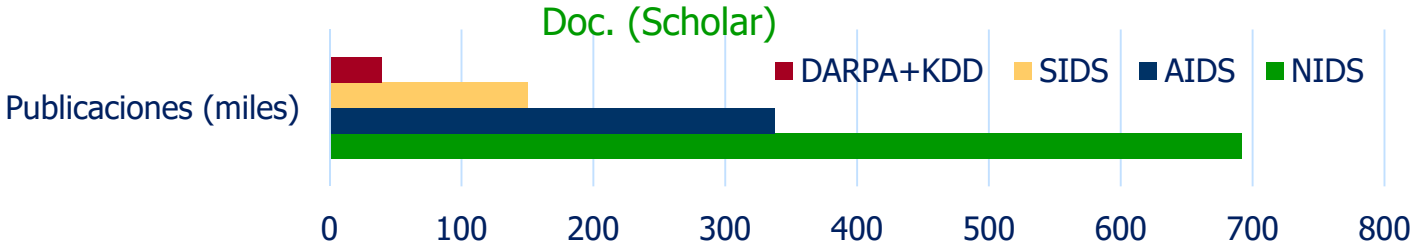
- Seguridad perimetral
  - Básicamente FW + IDS



© 2019 - Jesús E. Díaz Verdejo

Desarrollo/investigación en IDS

- Problema no resuelto
  - Primeros trabajos: Denning 86
- Gran volumen de actividad investigadora
  - Especialmente en detección de anomalías / híbrida



- Soluciones **desplegadas** basadas mayoritariamente en **firmas**

radware
http://www.ndm.net/ips/solutions/radware

Check Point SOFTWARE TECHNOLOGIES LTD.
http://www.ndm.net/ips/solutions/check-point

SOURCEfire
http://www.ndm.net/ips/solutions/sourcefire

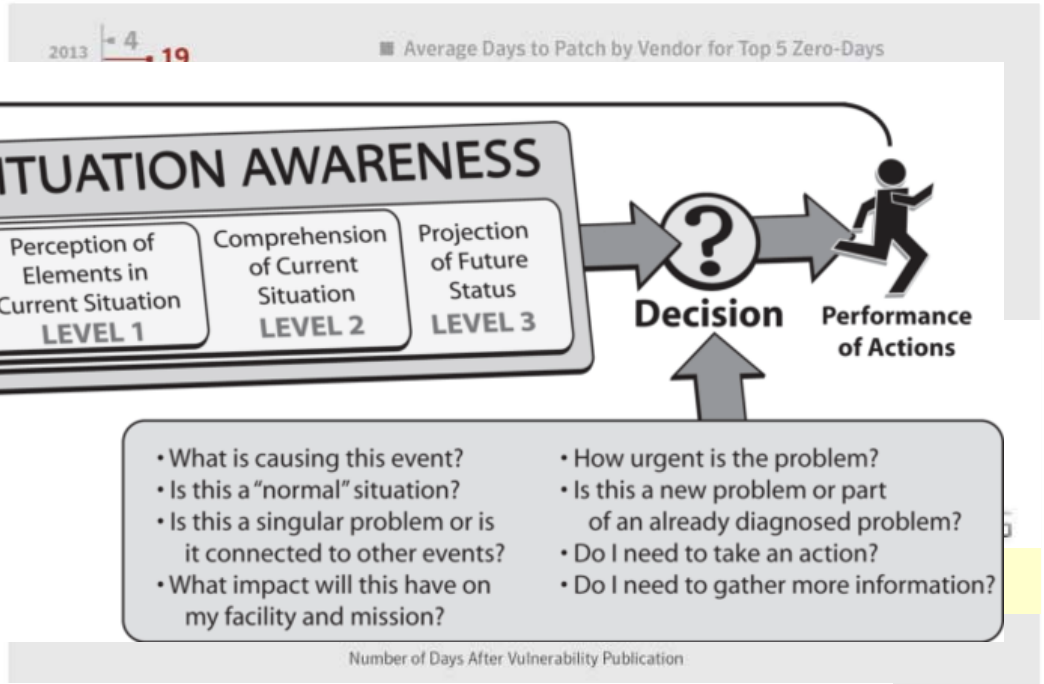
CISCO
SNORT

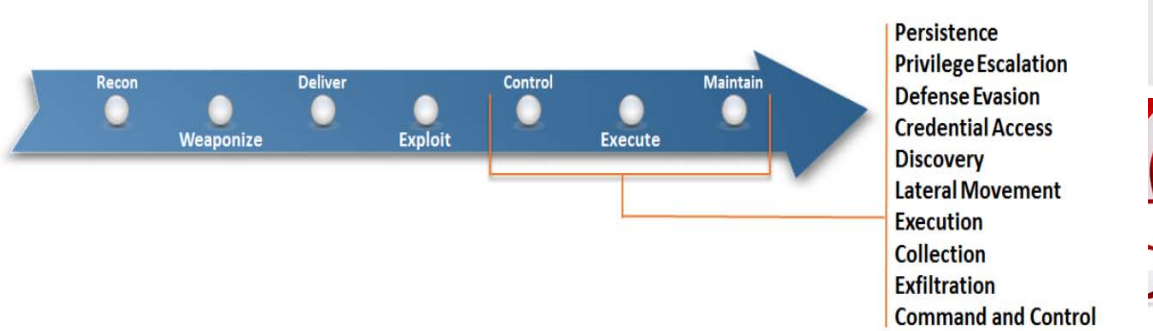
© 2019 - Jesús E. Díaz Verdejo

Desarrollo/investigación en IDS

- Insufic
  - (Firr
  - (Firr
  - Atac
  - APT
  - Opa

  ETPRO CURI
- Se nec
  técnica
  eventos
  - > 300k
    - Exce
  - **¿Por q**





© 2019 - Jesús E. Díaz Verdejo

Esquema

- Introducción
- Sistemas de detección de intrusiones
- Desarrollo de un IDS
- Deficiencias típicas
- Desafíos y futuro

© 2019 - Jesús E. Díaz Verdejo

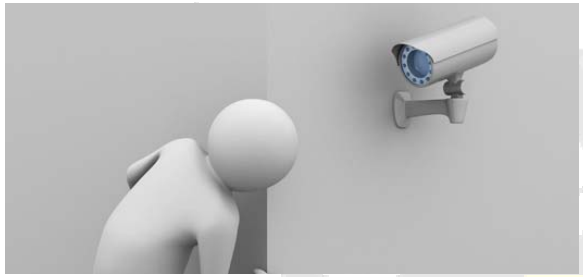
2. Sistemas de detección de intrusos

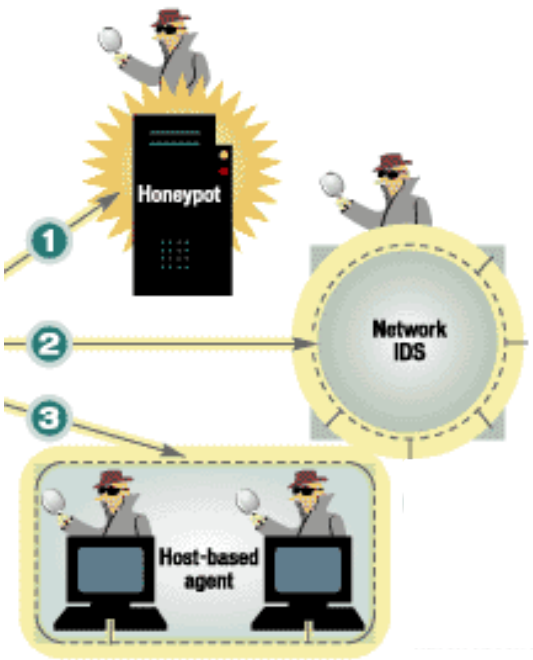
- Metodología básica
  - **Monitorizar eventos y analizarlos mediante los métodos apropiados**
- Elemento (fundamental) de NSM
  - Objetivo: generar **alertas** de intrusión
  - IPS: (y responder)
- Características:
  - Defensa centrada en vulnerabilidades
  - Foco en detección
  - Mayoritariamente basado en firmas

© 2019 - Jesús E. Díaz Verdejo

2. Sistemas de detección de intrusos

- **Detección de intrusiones**:
  - Proceso de monitorizar los eventos en un equipo o red en busca de signos de intrusión
  - Arte de detectar actividad inadecuada, incorrecta o anómala
- **Sistema de detección de intrusiones**
  - Autoexplicativo
    - Un administrador examinando trazas del sistema es un IDS
  - Combinación de software y/o hardware que intenta realizar la detección de intrusiones
  - ***¿Intrusiones?***





© 2019 - Jesús E. Díaz Verdejo

¿Ataques?

- **¿Qué es un ataque?**
  - Concepto bastante difuso
  - Se necesita establecer
    - **Servicios de seguridad**
      - Propiedades relevantes a la información
    - **Política de seguridad**
      - Conjunto de normas (reglas / leyes) que determinan qué está permitido o no en un sistema
  - **Violación de la política de seguridad**
    - ¡Clasificación de evento como ataque/no ataque dependiente del sistema!

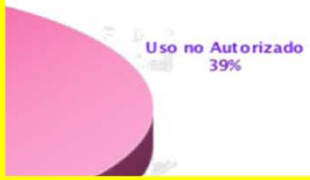
El IDS será específico para cada sistema

© 2019 - Jesús E. Díaz Verdejo

Tipos de ataques

- Di s

¿Qué técnica/s permite/n detectar qué tipos de ataques?



DoS/DD
2%

¿Qué eventos monitorizar?

- Ataques basados en ev
eventos
  - Ej. buffer overflow vs. DDoS

© 2019 - Jesús E. Díaz Verdejo

Indicadores de compromiso

- Elementos básicos para la detección
- **Indicador de compromiso (IOC):** Información que puede usarse para describir objetivamente una intrusión
  - Descripción independiente de la plataforma
  - Firmas como implementaciones de IOC
  - Una firma puede incluir uno o más IOC
- **Tipos**:
  - De red / de ordenador
    - Ejemplos
  - Estáticos / variables

De red	De ordenador
Dirección IPv4 / IPv6 Hash certificado X509 Nombre de dominio Cadena de texto Protocolo Nombre de archivo URL	Clave de registro Nombre de archivo Cadena de texto Nombre de proceso Mutex Hash de archivo Cuenta de usuario Path

© 2019 - Jesús E. Díaz Verdejo

IOC

Top 15 IOC

1. Tráfico saliente inusual
2. Anomalías en la actividad de cuentas con privilegios

Top Countries by Number of Attacks - Last 7 Days

Country	Block Count
China	13.094
Canada	4
France	2
United Kingdom	1
Japan	1

How are these categorized?

Total Attacks Blocked: Wordfence Network

24 Hours | 30 Days

Total Attacks

10.000.000
9.000.000
8.000.000

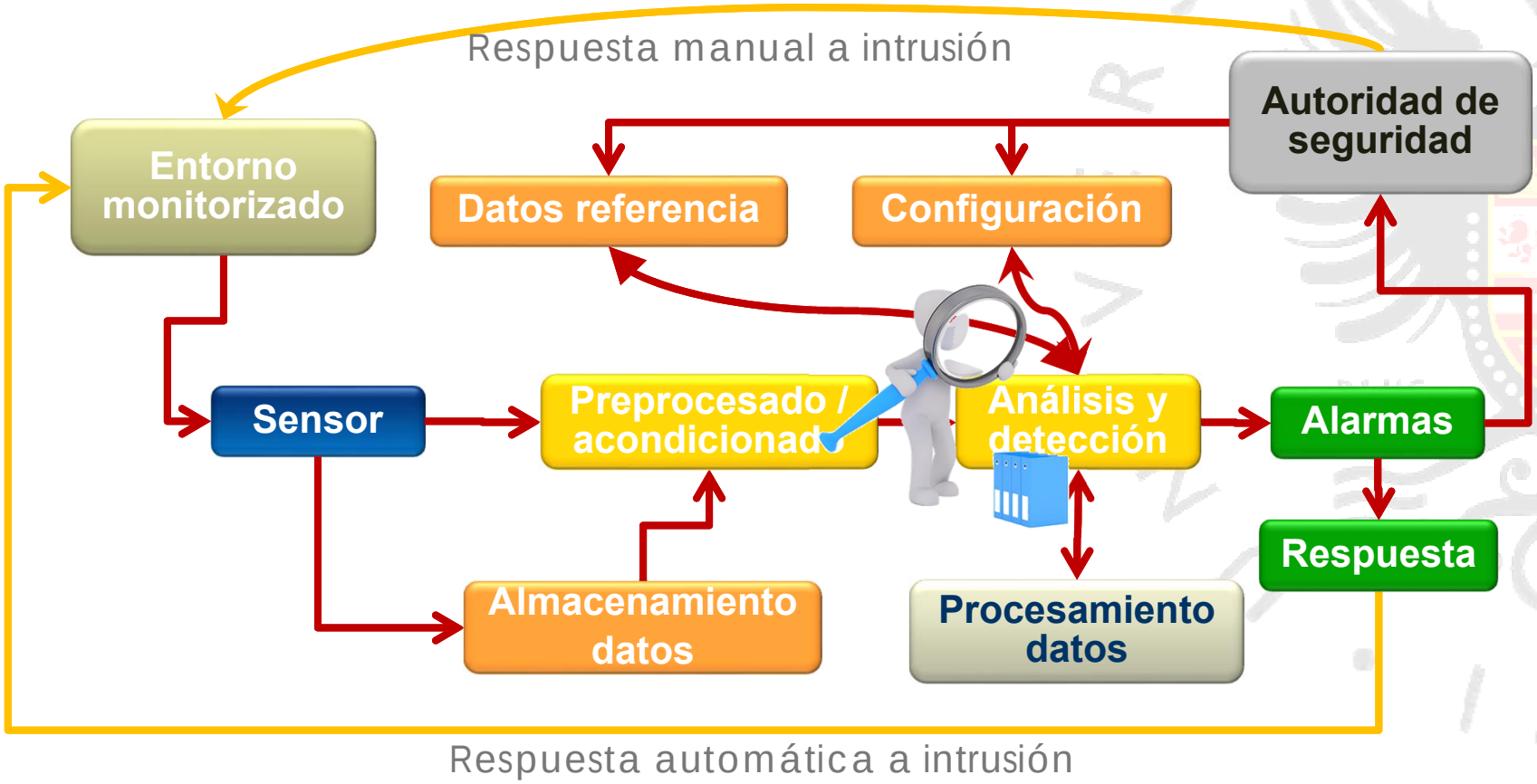
¿Qué eventos monitorizar?
(otra vez)

12. • Cambios en los perfiles movile
13. • Grandes volúmenes de datos e
14. • Tráfico web "no humano"
15. • Signos de DDoS

© 2019 - Jesús E. Díaz Verdejo

Elementos de un IDS

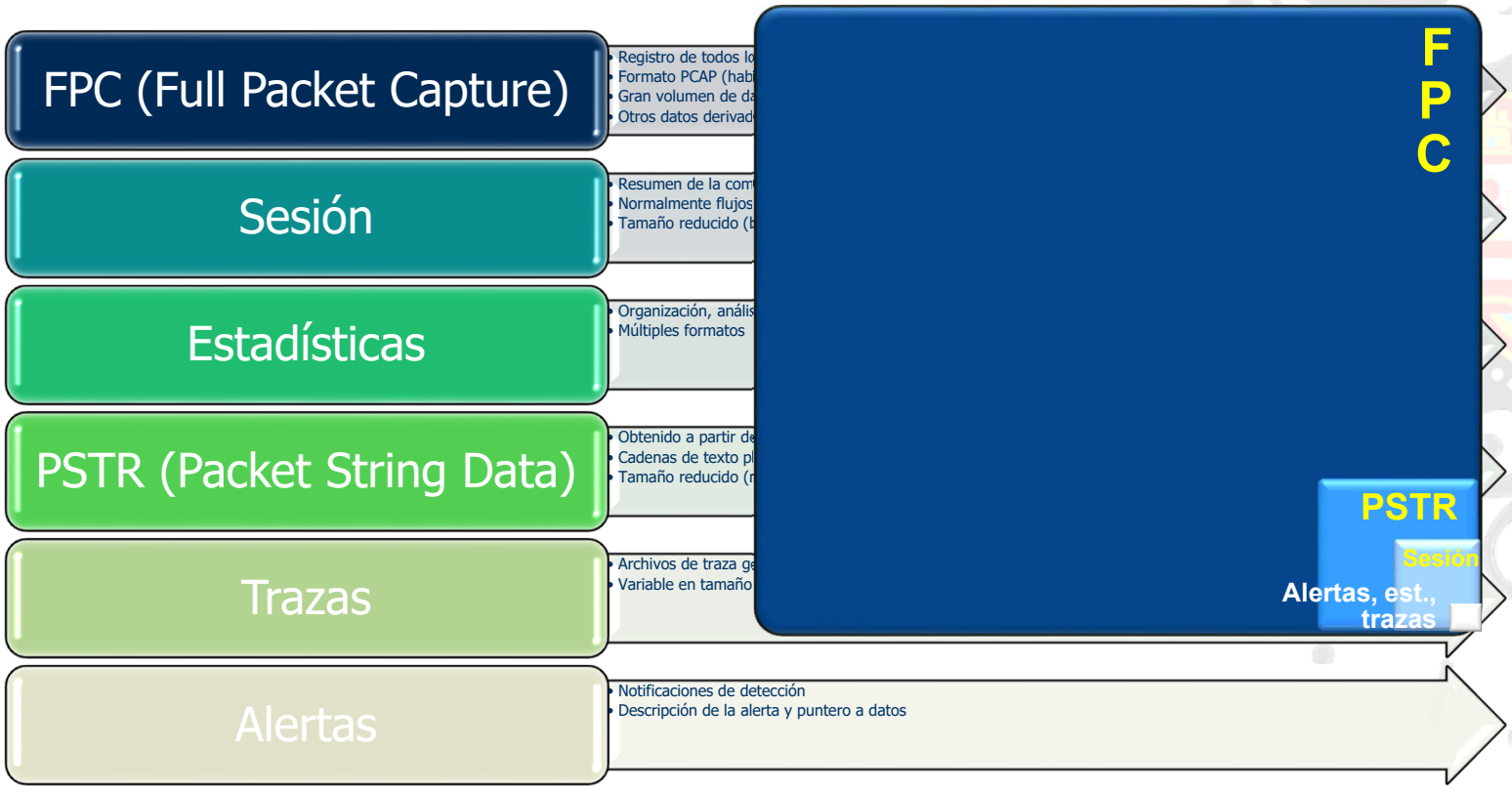
- Esquema práctico, con respuesta [Axelson,98], revisado



© 2019 - Jesús E. Díaz Verdejo

Elementos de diseño

- Decisiones/ técnicas para:
  - Eventos a monitorizar



© 2019 - Jesús E. Díaz Verdejo

Elementos de diseño

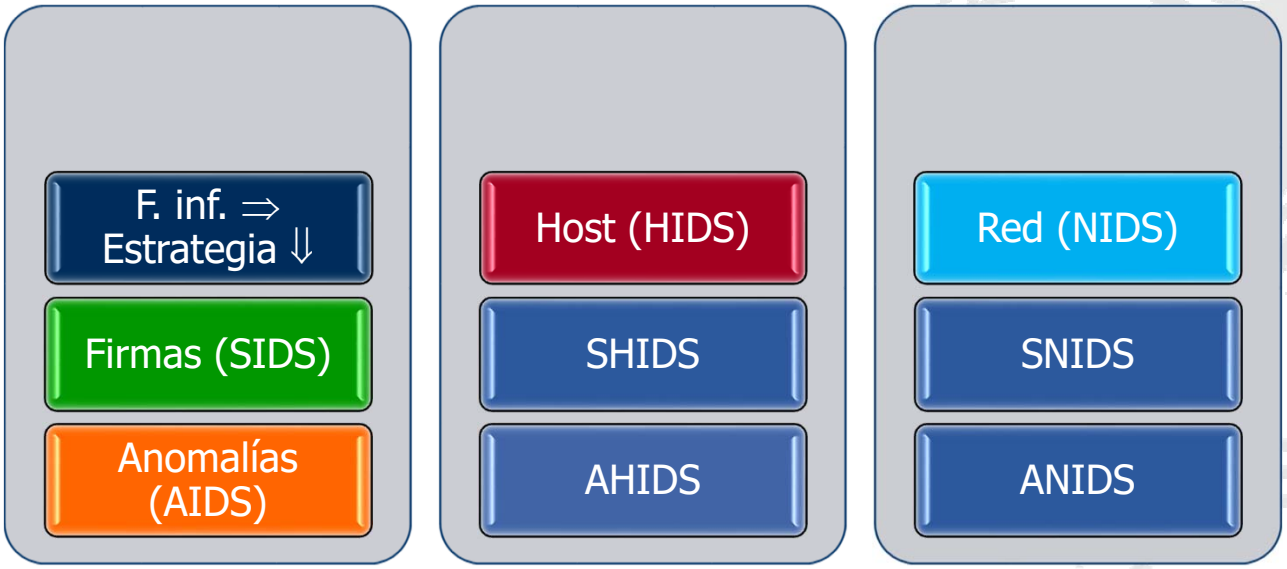
- Decisi
  - Eve
  - Pre

- Técnica de detección

© 2019 - Jesús E. Díaz Verdejo

Tipos IDS

- Múltiples clasificaciones posibles/propuestas
- Dos criterios (básicos):



© 2019 - Jesús E. Díaz Verdejo

IDS basados en firmas

- También denominados "***misuse***"-IDS
- Operación: búsqueda de actividad correspondiente a **ataques conocidos**
  - Basado en conocimiento extensivo de patrones asociados con ataques conocidos (**firmas**)

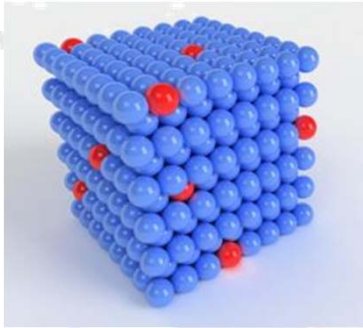
- Ej. Directory Traversal Vulnerability
  - *Solaris Sadmind/IIS worm (2001)*
    - Permite a peticiones HTTP GET cambiar directorio raíz con "`../../`".
    - Permite copiar `cmd.exe` al directorio de scri
    - Consecución control a nivel de administrado

Posibles firmas:
../..
CMD.exe+root.exe

```
GET/ scripts/../../winnt/system32/cmd.exe /c+
copy+\wint\system32\CMD.exe+root.exe
```

© 2019 - Jesús E. Díaz Verdejo

IDS basados en anomalías



- Operación: búsqueda de desviaciones respecto del comportamiento normal del sistema (**anomalía**)
  - ***Hipótesis de sospecha***: cualquier comportamiento sospechoso (diferente) es un ataque
    - Razonamiento implícito: explotar una vulnerabilidad implica un uso anormal del sistema
      - ⇒ los ataques pueden detectarse a partir de las anomalías (*outlayers*?)
  - Basado en el uso de **modelos implícitos o explícitos** del comportamiento normal
  - Aproximaciones: múltiples aproximaciones incluyendo técnicas estadísticas, estocásticas, sistemas expertos, minería de datos, redes neuronales, etc.
  - **"Closed world"**

© 2019 - Jesús E. Díaz Verdejo

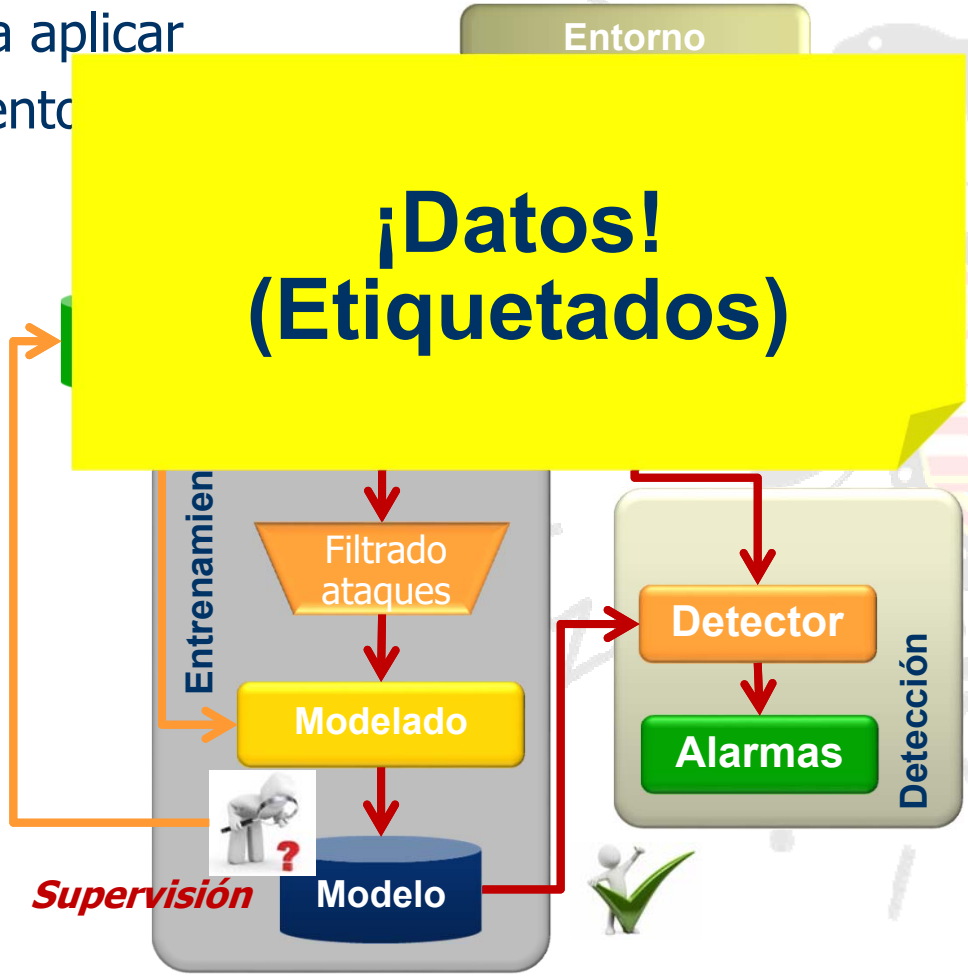
Esquema

- Introducción
- Sistemas de detección de intrusiones
- **Desarrollo de un IDS**
- Deficiencias típicas
- Desafíos y futuro

© 2019 - Jesús E. Díaz Verdejo

3. Desarrollo de un IDS

- Arquitectura y técnicas a aplicar
- Problema de entrenamiento y validación
- **Entrenamiento**
  - Estimación de los modelos de normalidad / firmas
  - Disposición de conjuntos suficientes de datos (bases de datos)
    - Tanto de ataque como normal (habitualmente)
    - **¡Etiquetado!**



© 2019 - Jesús E. Díaz Verdejo

3. Desarrollo de un IDS

Evaluación

- ¿Funciona?
- ¿Cuál es el rendimiento?
  - Comparación de IDS
  - Rendimiento absoluto v
    - Reproducible, abierto,
    - No necesariamente rep
      abierto y libre de sesg

Medidas de rendimiento

© 2019 - Jesús E. Díaz Verdejo

Medidas de rendimiento

- Clasificaciones posibles:

- Dos situaciones de error
- ¿Objetivo del detector?

© 2019 - Jesús E. Díaz Verdejo

Medidas de rendimiento

- Comportamiento deseable:
  - Pocos falsos positivos (idealmente, 0
    - Alto número de alarmas erróneas
      - El IDS se vuelve inútil
  - Ni
    -
  - Es
- Objet
  - D
    la
    re

SGUIL-0.9.0 - Connected To localhost

File Query Reports Sound: Off ServerName: localhost UserName: sar UserID: 2 2018-04-07 18:47:58 GMT

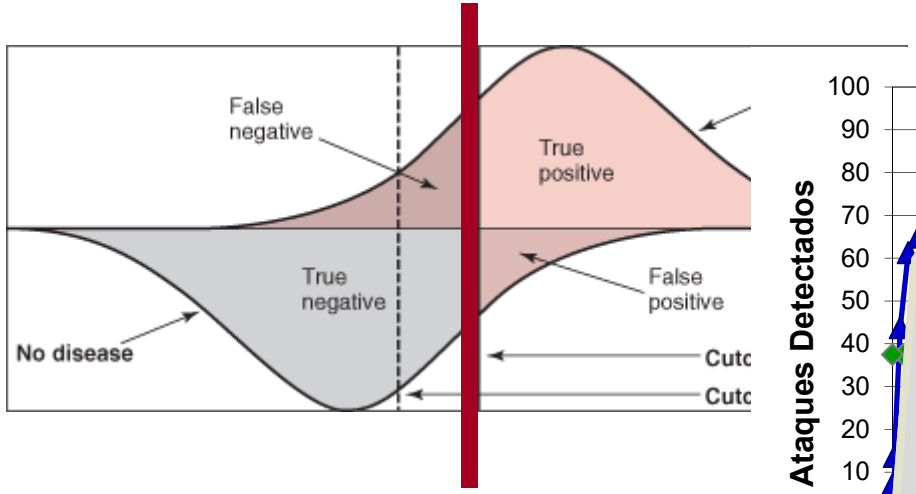
RealTime Events | Escalated Events

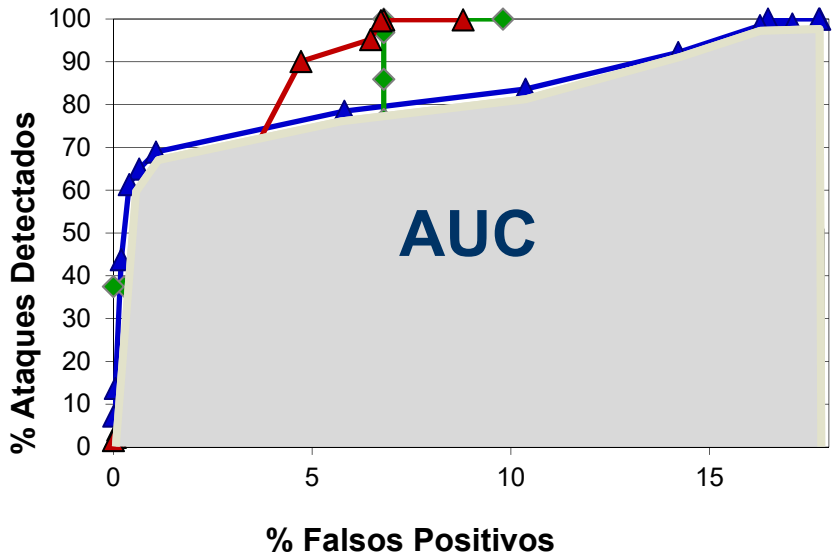
ST	CNT	Sensor	Alert ID	Date/Time	SrcIP	SPort	Dst IP	DPort	Pr	Event Message
RT	1	sar-Virtua...	8.1	2018-04-07 18:42:42	192.168.1.102	653	192.168.1.102	111	17	PADS New Asset - unknown @sunrpc
RT	1	sar-Virtua...	12.3	2018-04-07 18:42:42	192.168.1.102	61216	192.168.1.102	23	6	PADS New Asset - ftp ProFTPD Server 1.2.2rc3
RT	1	sar-Virtua...	8.2	2018-04-07 18:42:42	192.168.1.102	2243	192.168.1.102	21	6	PADS New Asset - ftp WU-FTPD Server 2.6.0(1)
RT	1	sar-Virtua...	8.3	2018-04-07 18:42:42	192.168.1.102	61216	192.168.1.102	23	6	PADS New Asset - ftp ProFTPD Server 1.2.2rc3
RT	1	sar-Virtua...	7.1	2018-04-07 18:42:42	210.114.220.46	653	192.168.1.102	111	17	GPL RPC portmap status request UDP
RT	1	sar-Virtua...	11.1	2018-04-07 18:42:42	210.114.220.46	653	192.168.1.102	111	17	GPL RPC portmap status request UDP
RT	1	sar-Virtua...	7.2	2018-04-07 18:42:42	210.114.220.46	654	192.168.1.102	919	17	GPL RPC STATD UDP stat mon_name format string exploit attempt
RT	2	sar-Virtua...	7.3	2018-04-07 18:42:42	192.168.1.102	23	217.156.93.166	61200	6	GPL TELNET Bad Login
RT	1	sar-Virtua...	7.5	2018-04-07 18:42:42	192.168.1.102	21	207.35.251.172	2243	6	ET POLICY FTP Login Successful
RT	37	sar-Virtua...	7.6	2018-04-07 18:42:42	207.35.251.172	2243	192.168.1.102	21	6	GPL FTP SITE EXEC attempt
RT	1	sar-Virtua...	11.2	2018-04-07 18:42:42	210.114.220.46	654	192.168.1.102	919	17	GPL RPC STATD UDP stat mon_name format string exploit attempt
RT	2	sar-Virtua...	11.3	2018-04-07 18:42:42	192.168.1.102	23	217.156.93.166	61200	6	GPL TELNET Bad Login
RT	36	sar-Virtua...	7.7	2018-04-07 18:42:42	207.35.251.172	2243	192.168.1.102	21	6	GPL FTP SITE overflow attempt

© 2019 - Jesús E. Díaz Verdejo

Medidas de rendimiento

- Determinación del rendimiento
  - Selección del punto de operación a partir de curvas ROC





- Otras figuras de mérito:
  - Rendimiento (núm. de eventos procesados)
  - Retardo en la detección/respuesta

© 2019 - Jesús E. Díaz Verdejo

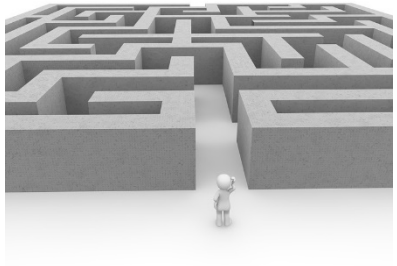
Elementos de diseño

- "Ingredientes"
  - El IDS será específico para cada sistema
  - ¿Qué eventos monitorizar?
  - Preprocesado
  - ¿Qué técnica/s permite/n detectar qué tipos de ataques?
  - Técnica de detección
  - ¡Datos! (Etiquetados)
  - Medidas de rendimiento

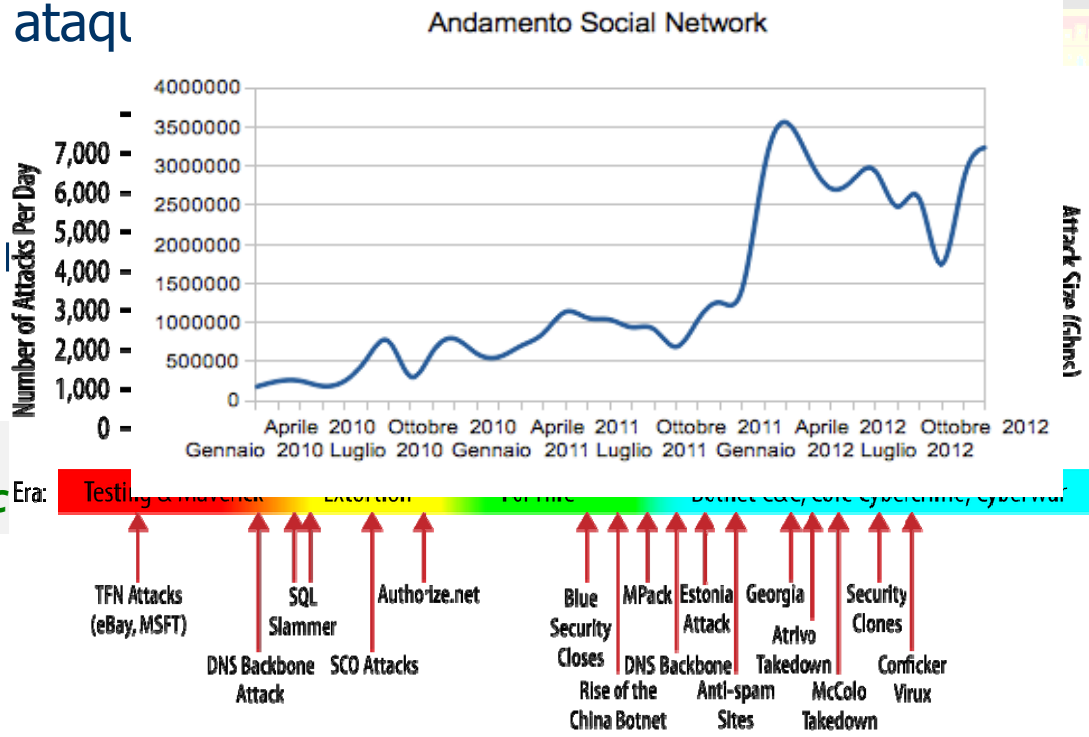
© 2019 - Jesús E. Díaz Verdejo

Elementos de diseño

- Otros ingredientes
  - Evolución de los sistemas
    - Problema de "data shift"
  - Evolución de los ataqu
  - Entorno adversar
    - Evasión



GET/
/c+ c



© 2019 - Jesús E. Díaz Verdejo

Esquema

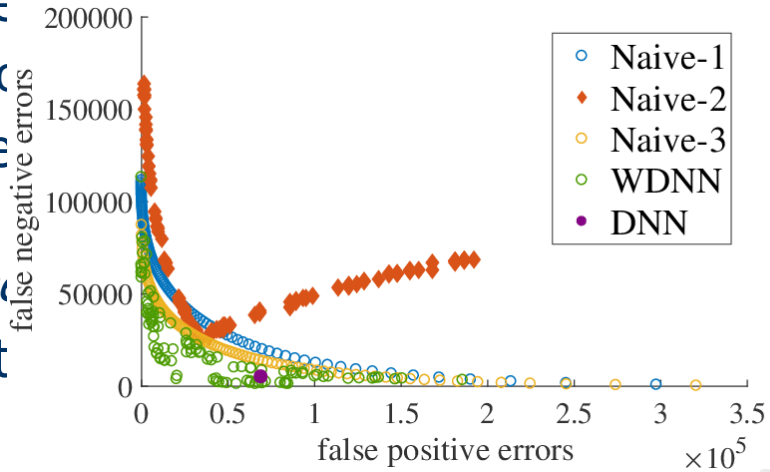
- Introducción
- Sistemas de detección de intrusiones
- Desarrollo de un IDS
- Deficiencias típicas
- Desafíos y futuro

© 2019 - Jesús E. Díaz Verdejo

4. Deficiencias típicas

Estado del arte

- Análisis deficie
  - No es un pro
- Aplicación "cie campos)
  - P.e. *machine*
- Es fácil subest



- Múltiples contribuciones se limitan a mostrar que su propuesta mejora los resultados de otra sobre el mismo conjunto de datos
  - Falta de generalidad
  - Sobreajustes (sobre-entrenamiento)

© 2019 - Jesús E. Díaz Verdejo

4. Deficiencias típicas

Estado del arte

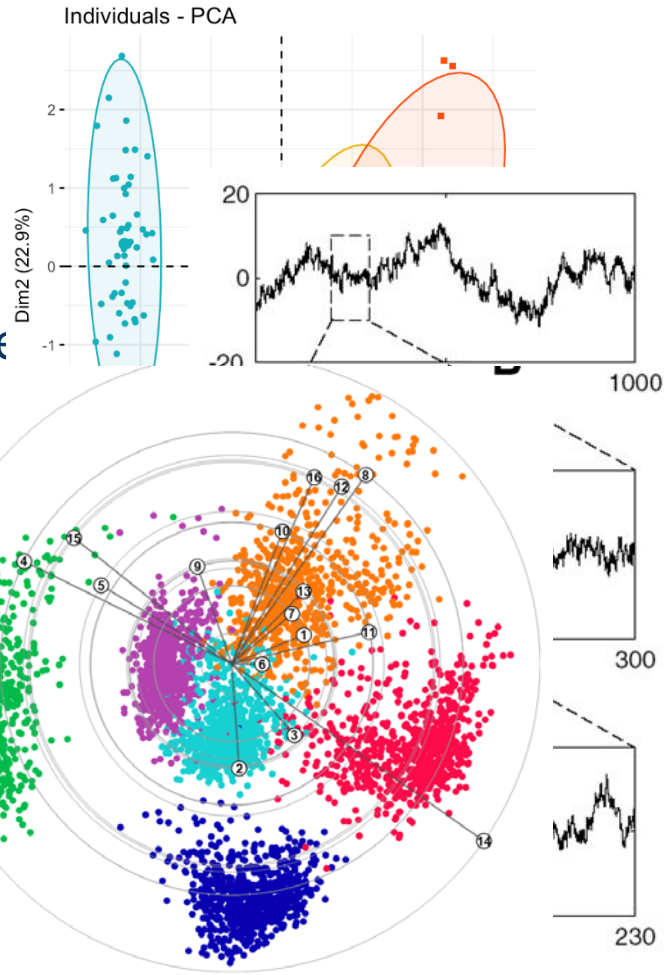
- Snort no es el estado del arte
  - A pesar de su gran penetración
- NIDS ya NO es:
  - Análisis basado en paquetes
  - Análisis basado en flujos
- Sintaxis vs. semántica

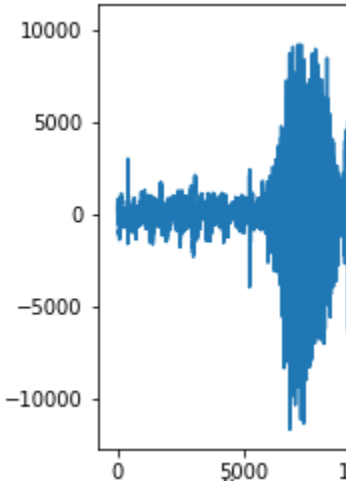
© 2019 - Jesús E. Díaz Verdejo

4. Deficiencias típicas

Preprocesado

- Datos en bruto
  - Aproximación típica Big Data / minería datos
  - Más datos ≠ mejores resultados (aume SNR)
- Parametrizacione
  - Más parámetros
  - Información disc asociada a mayo
- Granularidad (te
  - Hay evolución te mágnitudes
  - Tráfico autosimilar

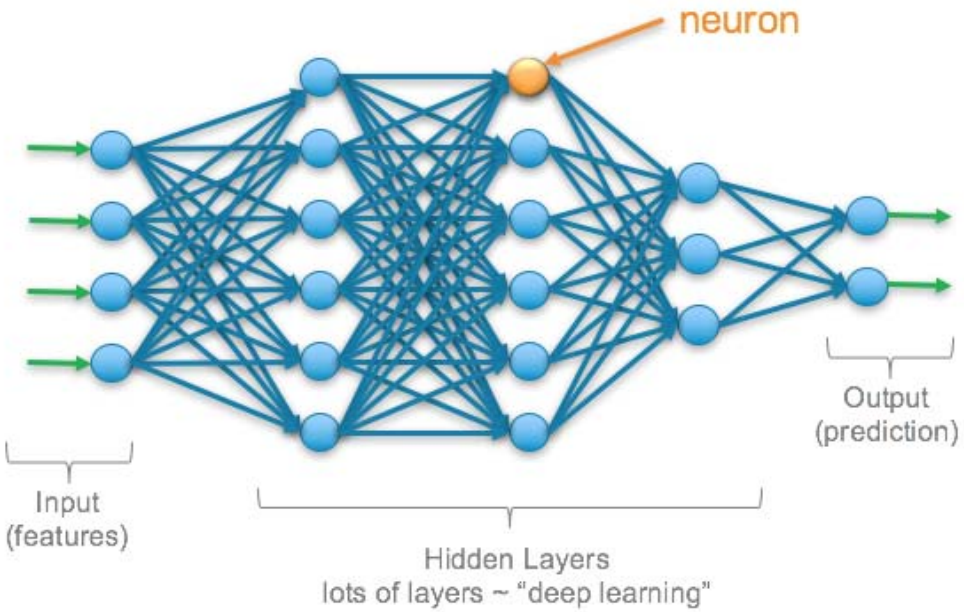


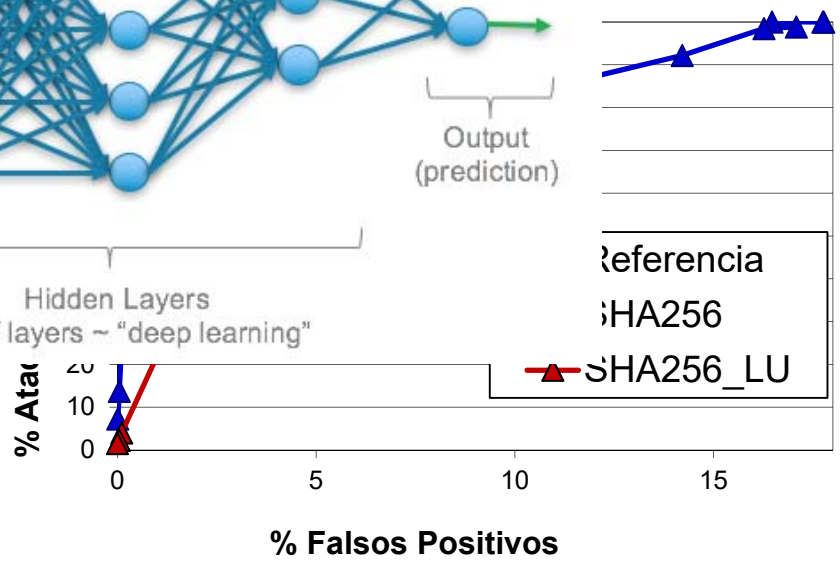


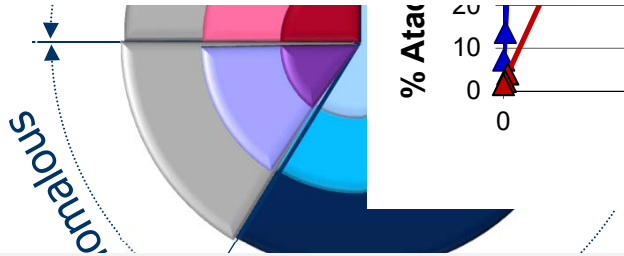
© 2019 - Jesús E. Díaz Verdejo

4. Deficiencias típicas

- **Entrena**
  - Particio
    - Misn
  - Medida
    - Uso
    - Punt
    - Sobr
      - F
  - Ajuste d
    - Núme
  - Capacid







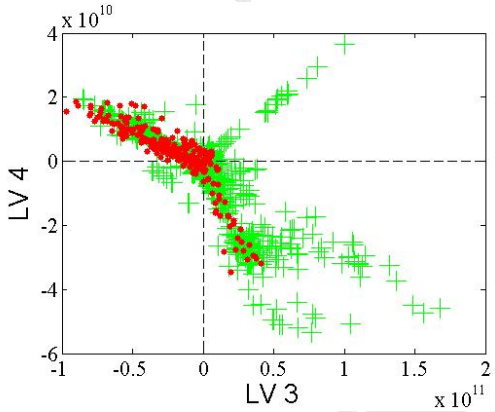
The Base-Rate Fallacy and its Implications for the Difficulty of Intrusion Detection, S. Axelsson, Proc. ACM CCS 1999

© 2019 - Jesús E. Díaz Verdejo

4. Deficiencias típicas

Datasets

- Falta de representatividad / escala
  - Datos propios
  - Simulaciones
- Obsolescencia
  - Evolución de los ataques en últimos años
  - Evolución de los sistemas en los últimos años
- Ocultismo
  - No se prop datasets
  - Presencia d
  - Artefactos
  - **Ground-truth**
- Tamaño
  - **Entrenamiento insuficiente**




```
NASA95: /htbin/wais.pl?wnet

BIBLIO:
/imce?app=ckeditor%7Csendto%40ckeditor_imceSend
To%7C&CKEditor=edit-body-und-0-
value&CKEditorFuncNum=1&langCode=es
```

© 2019 - Jesús E. Díaz Verdejo

DARPA / KDD99

- **DARPA'98 (y 99)**
  - Datos de Lincoln Labs (1998 y 1999)
  - Trazas de actividad, incluyendo ataques, de una hipotética base aérea
  - Una de las pocas bases de datos etiquetadas disponibles
- Problemas:
  - Sintética
    - Múltiples artefactos, p.e. TTL
  - Vieja
  - Sobreutilizada
- Se sigue utilizando todavía

- **KDD'99**
  - Versión parametrizada de DARPA'98
  - Propuesta para una competición de ML (KDD)
- Hereda los problemas de DARPA'98
- Parametrización inadecuada ("ciega")
- Aún más utilizada que DARPA'98

Testing Intrusion detection systems: A critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory, John McHugh, ACM Transactions on Information and System Security 3(4), 2000

© 2019 - Jesús E. Díaz Verdejo

Esquema

- Introducción
- Sistemas de detección de intrusiones
- Desarrollo de un IDS
- Deficiencias típicas
- Desafíos y futuro

© 2019 - Jesús E. Díaz Verdejo

Desafíos y futuro

- Que **funcione en condiciones reales**
  - Capacidad de generalización
  - Capacidad de adaptación
- **Retos técnicos** más relevantes
  - Elevado número de datos
  - Alta dimensionalidad de los datos
  - Naturaleza temporal
    - Los datos próximos en el tiempo suelen estar correlacionados
  - Distribución descompensada
    - La aguja en el pajar
  - Preprocesado de los datos

© 2019 - Jesús E. Díaz Verdejo

Desafíos y futuro

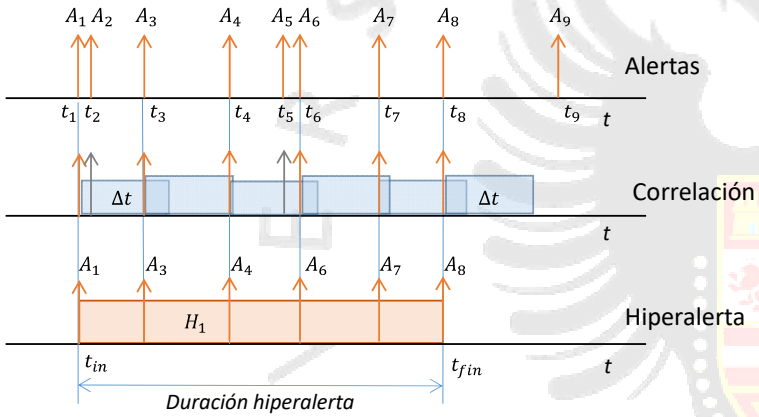
Problemas abiertos

- Baja tasa de eficiencia (problema de altas tasas de falsos positivos)
  - Se necesitan aproximaciones más estructuradas
  - Técnicas más eficientes
- Bajo "throughput" y alto coste
- Inexistencia de metodologías de evaluación válidas
- Bajo número de mecanismos de respuesta
  - Se requieren mecanismos más eficientes y más robustos
- Análisis de datos cifrados

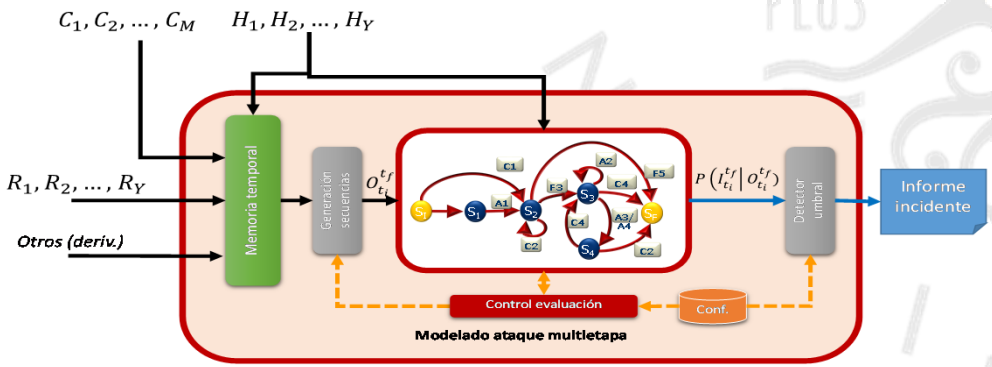
© 2019 - Jesús E. Díaz Verdejo

Desafíos y futuro

- Evolución hacia IDS híbridos distribuidos
- **Posprocesado**
  - Técnicas de correlación de eventos



- Modelado de (fases de) ataques
- Sensores adicionales



ATT&CK™
Adversarial Tactics, Techniques & Common Knowledge

© 2019 - Jesús E. Díaz Verdejo

Más información

- ***R. Sommer, V. Paxson;***
  Outside the Closed World: On Using Machine Learning For Network Intrusion Detection
  Proc. IEEE Symp. On Security and Privacy, 305-316, 2010.
- ***P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, E. Vazquez;***
  Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges
  Computers & Security, 28:18-28, 2009.
- ***M. Tavallaee, N. Stakhanova, A. Ghorbani,***
  Toward Credible Evaluation of Anomaly-Based Intrusion-Detection Methods
  IEEE Transactions on Systems, Man, And Cybernetics —Part C, 40(5):516-524, 2010.

© 2019 - Jesús E. Díaz Verdejo

Desarrollando sistemas de detección de intrusiones

Del laboratorio a su implantación

Facultad de Informática – Univ. Complutense de Madrid

Abril 2019

Jesús Esteban Díaz-Verdejo

Departamento de Teoría de la Señal, Telemática y Comunicaciones

E.T.S. Ingenierías Informática y Telecomunicación – Universidad de Granada

C/ Periodista Daniel Saucedo Aranda, s/n - 18071 – Granada (Spain)

Phone: +34-958 242304 - Email: jedv@ugr.es

© JEDV, 2005