Monitoring SIP Traffic Using Support Vector Machines

Mohamed Nassar, Radu State,
and Olivier Festor

Centre de Recherche INRIA Nancy - Grand Est
Villers-Lès-Nancy, France

Presenta: Juan Ramón Cayón Alcalde

upna Universidad Pública de Navarra Nafarroako Unibertsitate Publikoa

Contenidos

- [1] Amenazas
- [3] Sistema de Monitorización
- [5] 38 Parámetros
- [4] Support Vector Machine
- [5] Monitorizando SIP
- [3] Rendimiento y Precisión
- [6] Detección de Ataques
- [2] Conclusiones

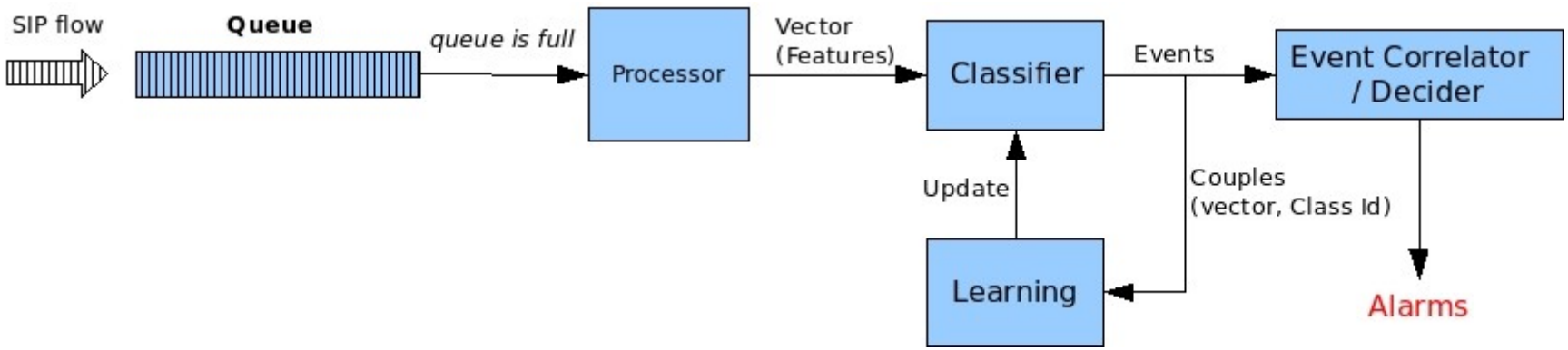
Amenazas

- Ataques por inundación:
  - “fáciles” de detectar
  - UDP, INVITE, INVITE/REGISTER
- Amenazas “sociales”
  - Difíciles de detectar (user dependant, mensajes bien formados, etc.)
  - Llamadas no solicitadas, **SPIT**, Vishing

Sistema de Monitorización

- Dividen la traza en intervalos fijos de tiempo
- Analizan 38 características en cada trozo
- Generan un vector de estadísticas con los resultados de los 38 parámetros
- El vector se pasa a un sistema de clasificación (SVM en este caso)
- El clasificador decide si el vector es “normal” o “anómalo”

Sistema de Monitorización



Sistema de Monitorización

- Basado en aprendizaje
- Alimentado con pares (vector, classId)
- Aprendizaje “al vuelo”
- Arquitectura modular:
  - <> Sistemas clasificación
  - <> Técnicas IA
- Tiempo real:
  - tpace < S/λ
    - tpace = tanalysis (proccessor)+ tmachine (classifier)

38 parámetros - G1

Group 1 - General Statistics		
1	Duration	Total time of the slice
2	NbReq	# of requests / Total # of messages
3	NbResp	# of responses / Total # of messages
4	NbSdp	# of messages carrying SDP / Total # of messages
5	AvInterReq	Average inter arrival of requests
6	AvInterResp	Average inter arrival of responses
7	AvInterSdp	Average inter arrival of messages carrying SDP bodies

- Permiten caracterizar el tráfico a rasgos generales (congestión,QoS)
- Los ataques por inundación provocan picos en todos estos parámetros

38 parámetros - G2

Group 2 - Call-ID Based Statistics		
8	NbSess	# of different Call-IDs
9	AvDuration	Average duration of a Call-ID
10	NbSenders	# of different senders / Total # of Call-IDs
11	NbReceivers	# of different receivers / Total # of Call-IDs
12	AvMsg	Average # of messages per Call-ID

- Equivalenetes al “modelo Erlang” en telefonía RTB
- Diálogos INVITE más largos y con mayor # de mensajes

38 parámetros - G3

Group 3 - Dialogs Final State Distribution		
13	NbNOTACALL	# of NOTACALL/ Total # of Call-ID
14	NbCALLSET	# of CALLSET/ Total # of Call-ID
15	NbCANCELED	# of CANCELED/ Total # of Call-ID
16	NbREJECTED	# of REJECTED/ Total # of Call-ID
17	NbINCALL	# of INCALL/ Total # of Call-ID
18	NbCOMPLETED	# of COMPLETED/ Total # of Call-ID
19	NbRESIDUE	# of RESIDUE/ Total # of Call-ID

- Estado final de una llamada en el momento del análisis (dentro de la partición)
- En condiciones normales predominan NOTACALL, COMPLETED y REJECTED

38 parámetros - G4

Group 4 - Requests Distribution		
20	NbInv	# of INVITE / Total # of requests
21	NbReg	# of REGISTER/ Total # of requests
22	NbBye	# of BYE/ Total # of requests
23	NbAck	# of ACK/ Total # of requests
24	NbCan	# of CANCEL/ Total # of requests
25	NbOpt	# of OPTIONS / Total # of requests
26	Nb Ref	# of REFER/ Total # of requests
27	NbSub	# of SUBSCRIBE/ Total # of requests
28	NbNot	# of NOTIFY/ Total # of requests
29	NbMes	# of MESSAGE/ Total # of requests
30	NbInf	# of INFO/ Total # of requests
31	NbPra	# of PRACK/ Total # of requests
32	NbUpd	# of UPDATE/ Total # of requests

- IM: SUBSCRIBE/NOTIFY (presencia) y MESSAGE (mensajes)
- VoIP-->PSTN: PRACK (fiabilidad respuestas provisionales)
- PSTN-->VoIP: INFO (tonos DTMF)

38 parámetros - G5

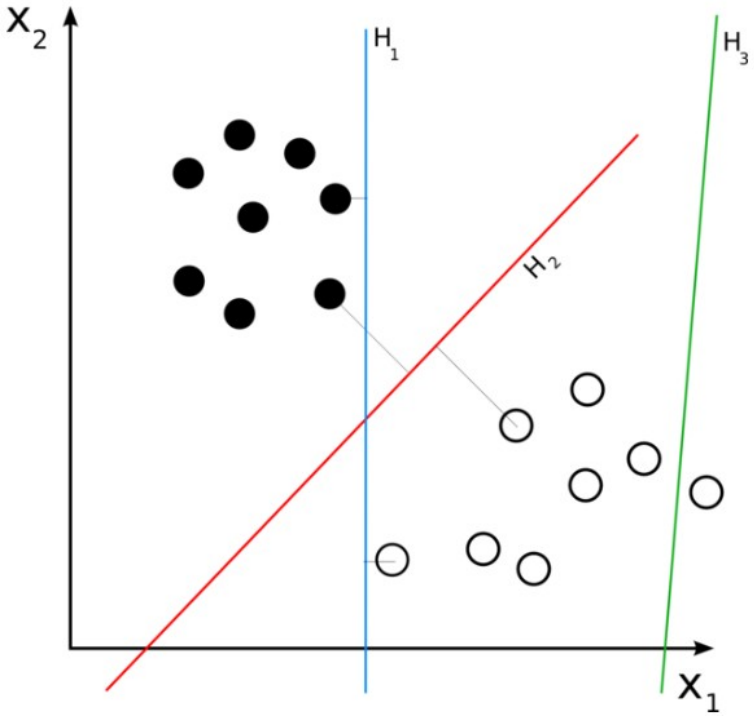
Group 5 - Responses Distribution		
33	Nb1xx	# of Informational responses / Total # of responses
34	Nb2xx	# of Success responses / Total # of responses
35	Nb3xx	# of Redirection responses / Total # of responses
36	Nb4xx	# of Client error responses / Total # of responses
37	Nb5xx	# of Server error responses / Total # of responses
38	Nb6xx	# of Global error responses / Total # of responses

- Alta tasa de errores --> algo estará pasando

Support Vector Machine

- Conjunto de métodos de aprendizaje supervisado, empleados en clasificación y regresión.
- Dado un espacio n-dimensional en el que existen dos conjuntos de vectores, SVM trata de construir un hiperplano que separe ambos conjuntos maximizando el margen entre ambos.
- A mayor margen, menor error en la clasificación.

Support Vector Machine



$$\text{Find } \vec{w} \text{ and } b \text{ to minimize } \frac{1}{2}\vec{w}.\vec{w}$$
$$\text{so that } y_l(\vec{w}.\vec{x_l}) + b \geq 1 \forall (\vec{x_l}, y_l) \in S$$

Método Lineal

$$\text{Find } \vec{w}, \text{ b and } \xi \text{ to minimize } \frac{1}{2}\vec{w}.\vec{w} + C\sum_l \xi_l$$
$$\text{so that } \begin{cases} y_l(\vec{w}.\vec{x_l}) + b \geq 1 - \xi_l, \forall (\vec{x_l}, y_l) \in S \\ \xi_l \geq 0, \forall l \end{cases}$$

C-SVM

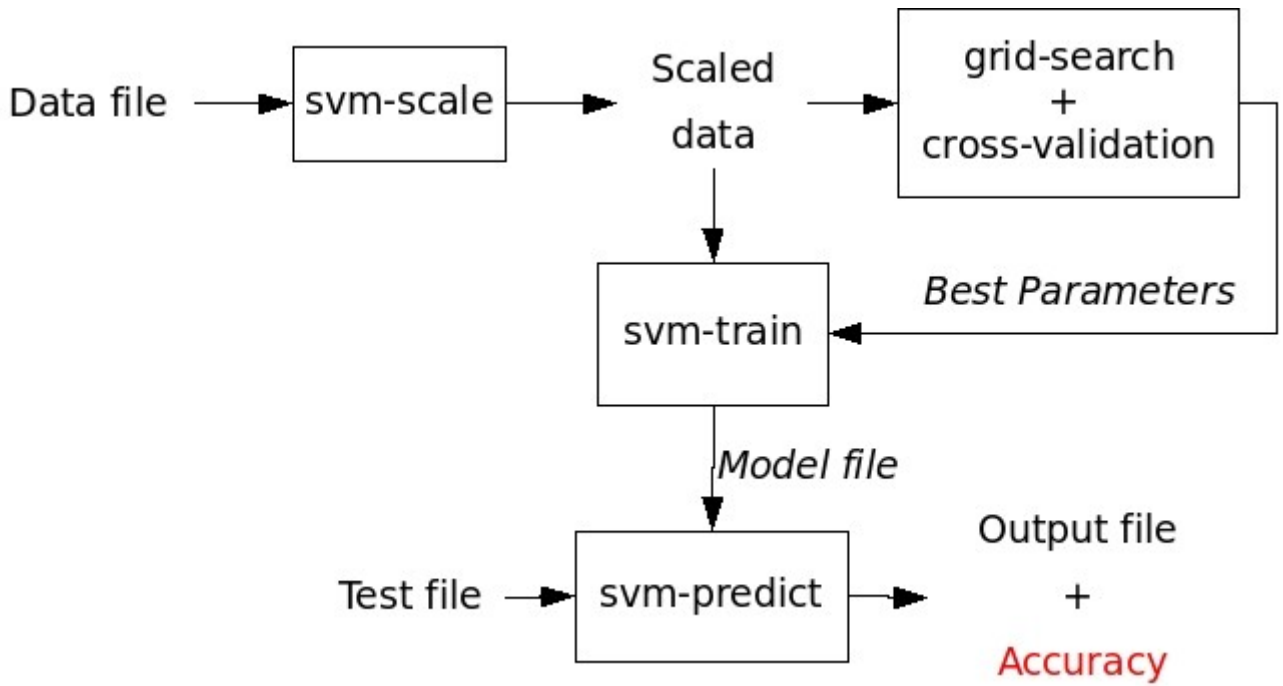
$$\text{linear } K_l(\vec{x}, \vec{z}) = \vec{x}.\vec{z}$$
$$\text{polynomial } K_d(\vec{x}, \vec{z}) = (\gamma \vec{x}.\vec{z} + r)^d \ , \gamma > 0$$
$$\text{radial basis function } k_{rbf}(\vec{x}, \vec{z}) = exp(-\gamma|\vec{x} - \vec{z}|^2) \text{ where } \gamma > 0$$
$$\text{sigmoid } k_s(\vec{x}, \vec{z}) = tanh(\gamma \vec{x}.\vec{z} + r), \ \gamma > 0 \text{ and } r < 0$$

Distintos Kernels

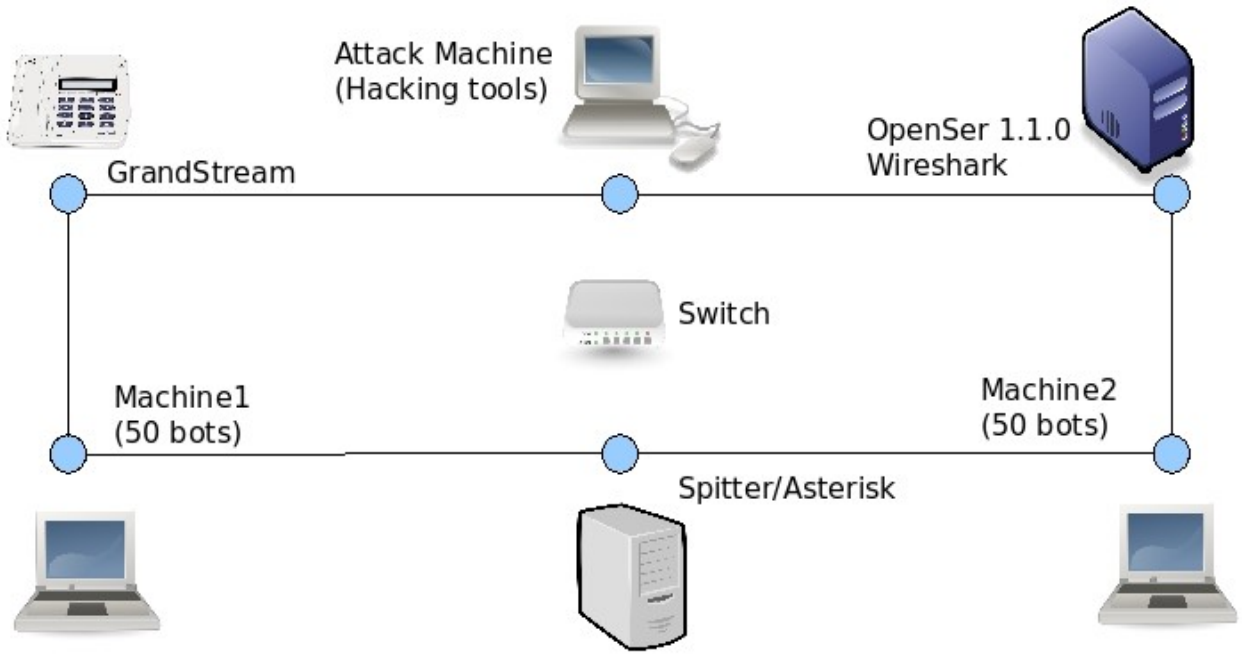
Support Vector Machine

- Implementada mediante librería LibSVM:
  - Algoritmos de clasificación y regresión
  - Multi-class classification
  - Estimadores de probabilidad
  - One-class training
  - Precisión de los tests:
    - Algoritmo de búsqueda grid sobre el espacionde parámteros SVM
    - Optimización de parámetros mediante validación cruzada

Support Vector Machine



Monitorizando SIP - Testbed

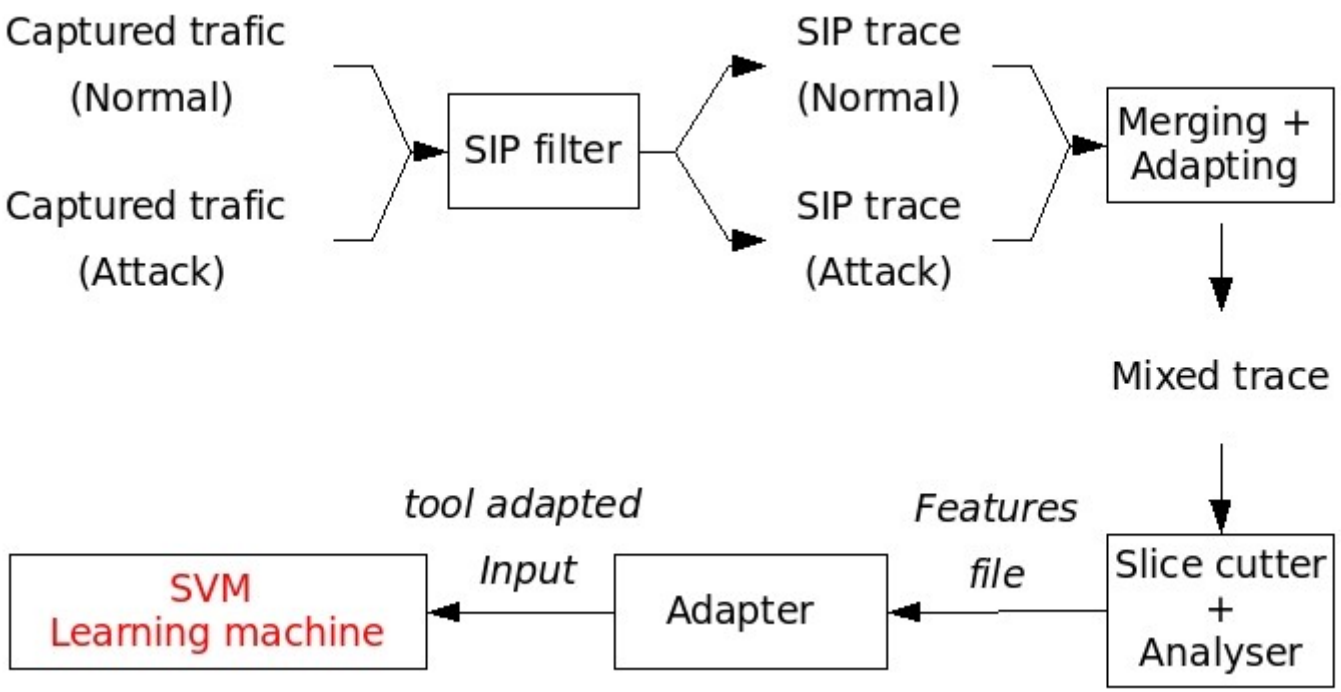


- Hacking tools: scanning, flooding, SPIT
- Bots: SIP UAs programables, controlados x canal IRC
- 1 único dominio
- Todos los UAs registrados en el servidor
- Las trazas se capturan en el servidor

Monitorizando SIP - Trazas

- Emplean 3 tiposde trazas:
  - Tráfico VoIP real “limpio” (captura de dos días de tráfico de un proveedor de servicio VoIP, sin ataques)
  - Tráfico de ataques generado en su maqueta
  - Traza mixta, mezclando tráfico de las trazas anteriores

Monitorizando SIP - Trazas



Monitorizando SIP - Tráfico Normal

- Métodos más abundantes: REGISTER y OPTIONS
- Métodos ausentes: MESSAGE, PRACK y UPDATE
- NOTIFY presenta estadísticas constantes en los distintos períodos del día (indica que los UAs están permanentemente conectados)
- INVITE, BYE y ACK presentan ratios casi constantes en todos los períodos con promedios #INVITE/#BYE = 2.15 #INVITE/#ACK = 0.92

Monitorizando SIP - Tráfico Normal

- Respuestas dominantes 2xx
- Respuestas 3xx-5xx-6xx apenas aparecen
- #INVITE/#1xx = 0.59
- Tiempos promedio de llegadas i/peticiones e i/respuestas constantes en todos los períodos (20ms)
- Tiempos promedio de llegadas i/peticiones con SDP entre 3s (horas valle) y 0,5s (horas punta)

Rendimiento y Precisión

Table 1. Coherence Test for two Successive Days

Day 1	06-10	10-14	14-18	18-22
Day 2	06-10	10-14	14-18	18-22
Accuracy(%)	55.91	53.72	52.83	56.90

Table 2. Coherence Test for Different Periods of the Same Day

Day 1	02-06	02-06	02-06	02-06	22-02
Day 1	06-10	10-14	14-18	18-22	22-02
Accuracy(%)	51.82	62.79	63.72	63.76	60.80

Rendimiento y Precisión

Table 3. Multi-Class SIP Traffic Data Set

Trace	Normal	DoS	KIF	Unknown
SIP pkts	57960	6076	2305	7033
Duration	8.6(min)	3.1(min)	50.9 (min)	83.7(day)

Table 4. Testing Results for Different Kernels

Kernel	Parameters	Accuracy(%)	Time(ms)
Linear	$C = 1$	99.79	196
Polynomial	$C = 1$; $\gamma = 1/38$; $r = 0; d = 3$	79.09	570
Sigmoid	$C = 1$; $\gamma = 1/38$; $r = 0$	93.83	994
RBF	$C = 1$; $\gamma = 1/38$	98.24	668
Linear	$C = 2$	99.83	157
RBF	$C = 2$; $\gamma = 0.5$	99.83	294

Rendimiento y Precisión

Table 5. Testing Results for Different Kernels

Window size	5	15	30	60	90	120	150
Accuracy (%)	95.4	99.32	99.30	99.67	99.63	100	100
Analysis Time (min)	1.12	2.40	2.56	4.31	6.39	7.42	8.51

Tamaño de la partición de la traza

Table 6. Results for Decreasing Size of Features Set

# of features	38	31	18	12	7
Accuracy (%)	99.30	99.39	98.90	98.65	98.22
Machine Time (s)	1.85	1.59	1.42	1.28	0.57

Detección de Ataques - Flooding

- Inviteflood hacking tool
- 5 ataques, de 1m cada uno, a diferentes tasas (#INVITE/sg)
- Cada ataque se inyecta en una traza de 2h de tráfico “limpio”
- Se fija el inicio del ataque a los 5m de empezar la traza de 2h
- Aprendizaje con 100 INVITE/sg

Detección de Ataques - Flooding

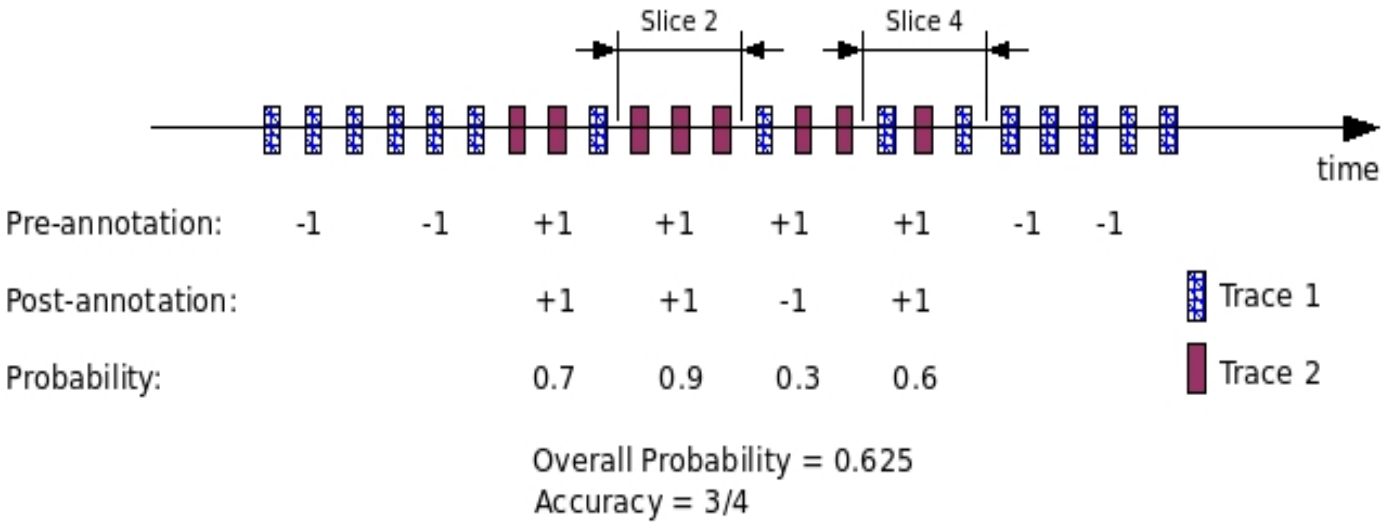


Table 7. Attack Estimation for Different Rates of Flooding

Flooding Rate (INVITE/s)	0.5	1	10	100	1000
Detection Accuracy-1 (%)	0	0	5.47	67.57	97.36
Detection Accuracy-2 (%)	0	1.48	30.13	88.82	98.24
Pr(Normal)	0.96	0.95	0.73	0.24	0.07
Pr(Attack)	0.04	0.05	0.27	0.76	0.93

Detección de Ataques - SPIT

- Spitter/Asterisk hacking tool
- Reacciones de los UAs (random):
  - Dejar sonar 1-6s y contestar
  - Contestar con “Ocupado”
  - Dejar sonar un rato e informar de Redirección
- 2 experimentos: SPIT parcial y SPIT total

Detección de Ataques - SPIT

- SPIT Parcial:
  - 100 números
  - Sólo 10 son UAs asociados al server
  - 10%hits (mundo real)
  - 4 ataques: 1, 10, **50**, 100 llamadas concurrentes
  - Misma inyección que en el flooding
- SPIT Total:
  - Todos los números son válidos (asociados)
  - El resto todo igual

Detección de Ataques - SPIT

Table 8. Detection of Partial SPIT in Four Mixed Traces With Different Intensities

# of Concurrent Calls	True Positives (%)	True Negatives (%)
RBF; C= 1; $\gamma = 1/38$; Training accuracy = 99.0249		
1	0 (0/3697)	100
10	1.30 (10/766)	
50	10.01 (62/619)	
100	18.31 (102/557)	
Linear ; C=1 ; Training accuracy = 99.0197		
1	0 (0/3697)	100
10	2.09 (16/766)	
50	10.66 (66/619)	
100	19.39 (108/557)	

Predicate	SPIT intensity
10 distributed positives in a 2 minutes period	Low
Multiple Series of 5 Successive Positives	Medium
Multiple Series of 10 Successive Positives	High

Reglas determinístas que mejoran las detecciones

Detección de Ataques - SPIT

Table 9. Detection of Full SPIT in Four Mixed Traces With Different Intensities

# of Concurrent calls	1	10	50	100
RBF; C= 1; $\gamma = 1/8$; Training accuracy = 98.9057				
True Positives	0.03 2/7015	3.05 15/492	12.18 85/698	23.41 184/786
True Negatives	100			

- Los resultados mejoran algo
- El nivel de anormalidad en el patrón de tráfico es idéntico en ambos casos (parcial y total)

En resumen...

- Se propone un sistema de monitorización online basado en SVM:
  - Se trocea la traza de señalización SIP
  - Se extrae un vector de datos (a partir del estudio de una serie de parámetros) que caracteriza a cada trozo
  - Se envía el vector a una SVM para su clasificación, basasándose en un modelo de aprendizaje
- Los resultados muestran:
  - El sistema es capaz de funcionar en tiempo real
  - Es capaz de detectar ataques de Inundación y SPIT, de manera precisa
  - La adición de determinadas reglas de correlación de eventos mejora la precisión

Líneas futuras...

- Redefinir y reordenar el conjunto de parámetros a estudio (mediante algoritmos de selección de parámetros)
- Detección de otros tipos de ataques

Monitoring SIP Traffic Using Support Vector Machines

Mohamed Nassar, Radu State,
and Olivier Festor

Centre de Recherche INRIA Nancy - Grand Est
Villers-Lès-Nancy, France

Presenta: Juan Ramón Cayón Alcalde