Monitoring SIP Traffic Using Support Vector Machines

Mohamed Nassar, Radu State, Olivier Festor

(nassar, state, festor)@loria.fr

MADYNES Team
INRIA, Nancy Grand Est

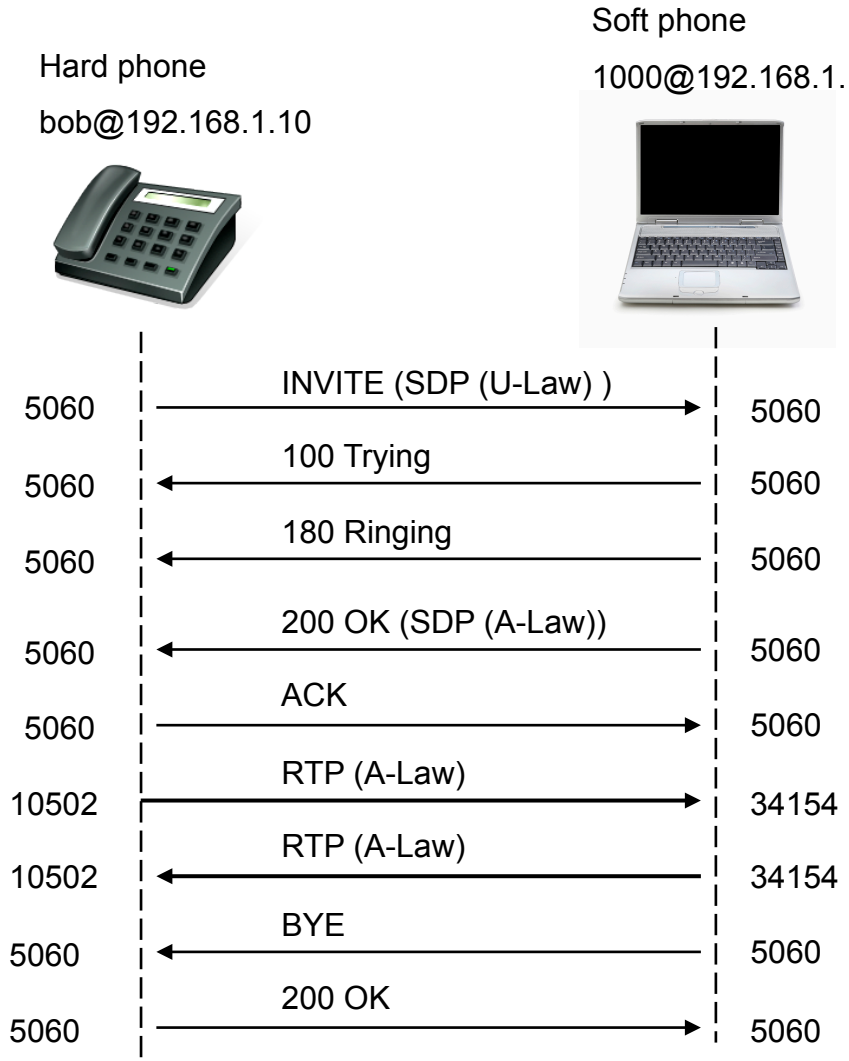
17 September 2008

Outline

- Introduction to SIP
- Threats
- Monitoring system
- Experiments
- Future works and Conclusion

SIP

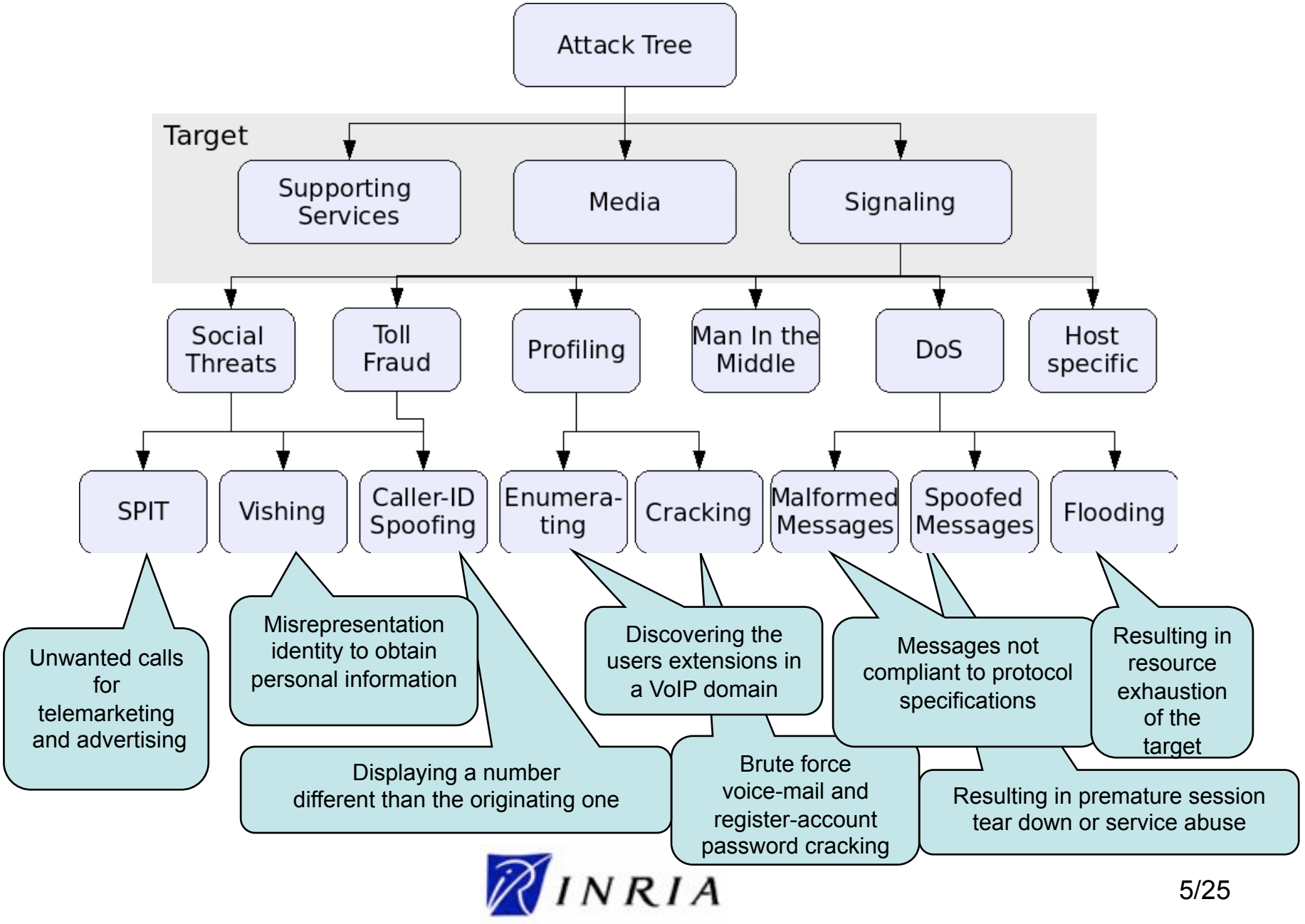
- SIP (Session Initiation Protocol - RFC 3261) Text-based like HTTP
- Request + response = transaction
- URI = *sip:user@host:port;parameters*



SIP Trapezoid

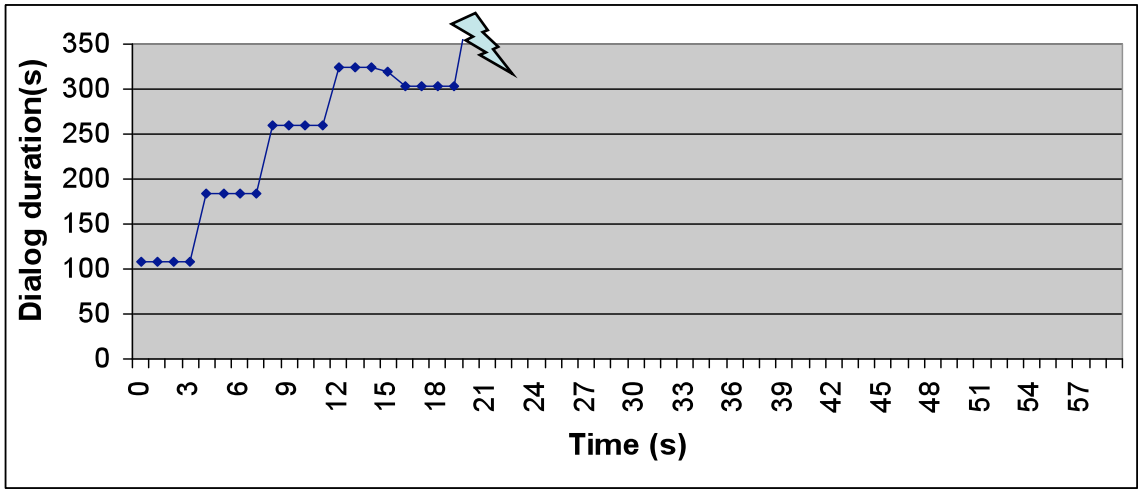

```
INVITE sip:Alice@berlin.org SIP/2.0
Via: SIP/2.0/UDP loria.nancy.org:5060;branch=z9hG4bKfw19b
Max-Forwards: 70
To: Alice <sip:Alice@berlin.org>
From: Bob <sip:Bob@nancy.org>;tag=76341
Call-ID: 123456789@loria.nancy.org
CSeq: 1 INVITE
Contact: <sip:Bob@nancy.org>
Content-Type: application/sdp

<SDP body not shown>
```

Threats in the VoIP domain



DoS

Using invalid destination domains with 100 Invite/second



- Flooding attacks target the signaling plane elements (e.g. proxy, gateway, etc.) with the objective to take them down or to limit their quality, reliability and availability

Strategy
Legitimate SIP messages
Malformed SIP messages
Invalid SIP messages
Spoofed SIP messages
CPU-based attacks targeting the authentication process

Destination
A valid URI in the target domain
A non existent URI in the target domain
A URI with an invalid domain or IP address
An invalid URI in another domain
A valid URI in another domain.

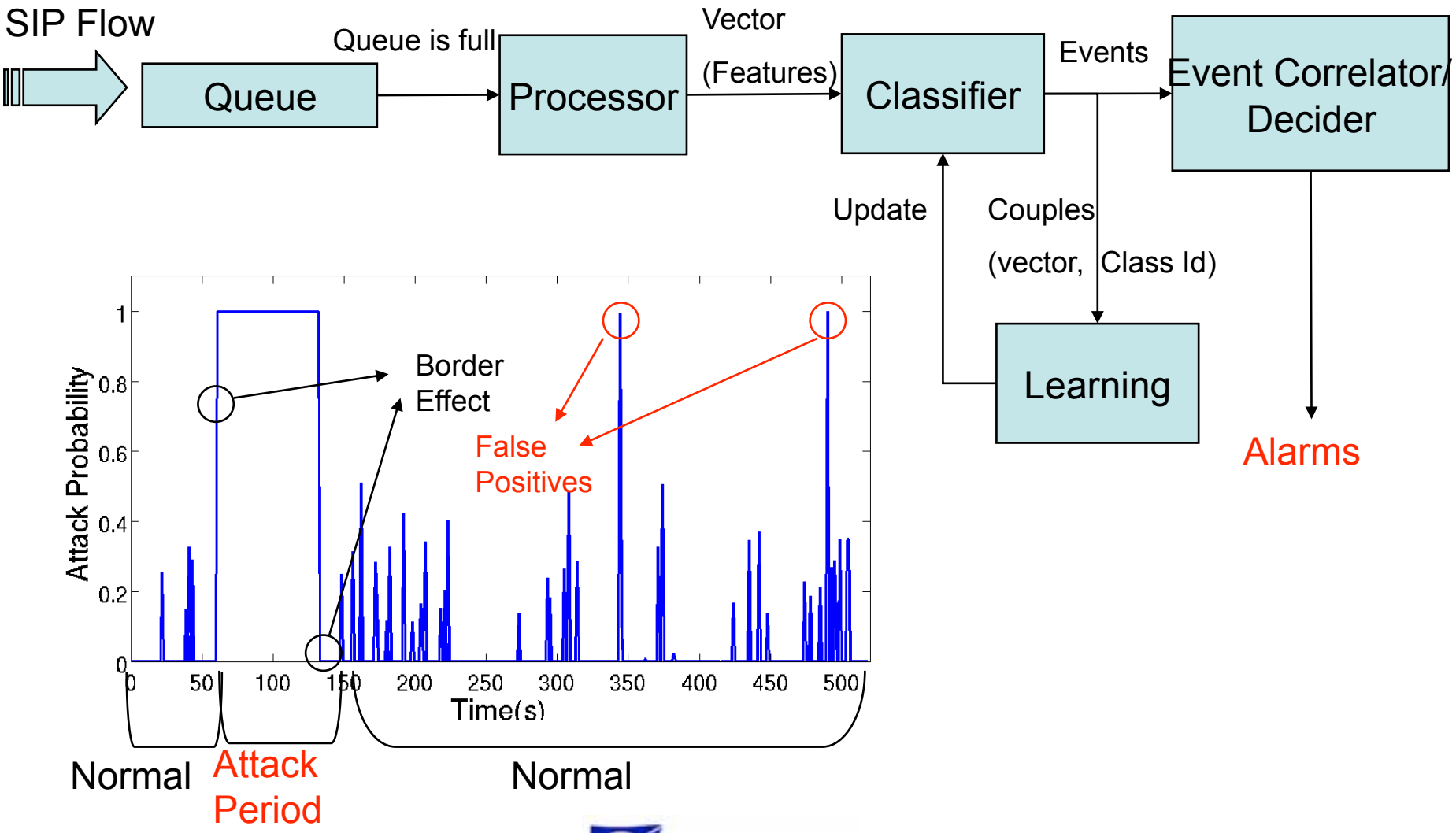
SPIT or SPam over Internet Telephony

- Like SPAM (cost-free) but more annoying (phone ringing all the day, interruption of work)
- Expected to become a severe issue with the large deployment of VoIP services
- SPIT transactions are technically correct
- We don’t know the content until the phone rings
- We need to be reachable
- SPAM filtering solutions are not directly applicable
- Current approaches: multi-level grey list, Turing tests, Trust management, VoIP SEAL from NEC, VoIP SPAM detector from University of North Texas

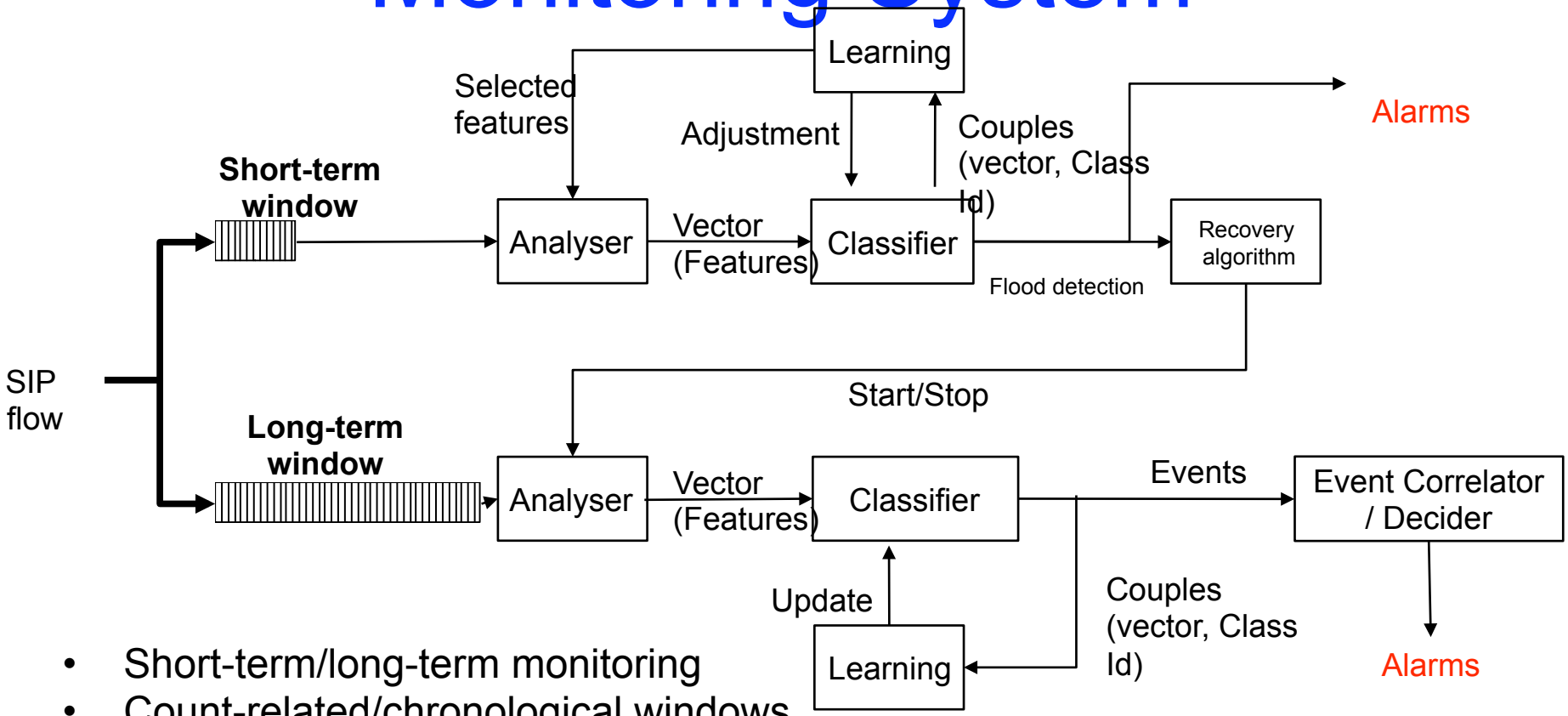
**From winnipeg.ca*

INRIA

Monitoring Approach

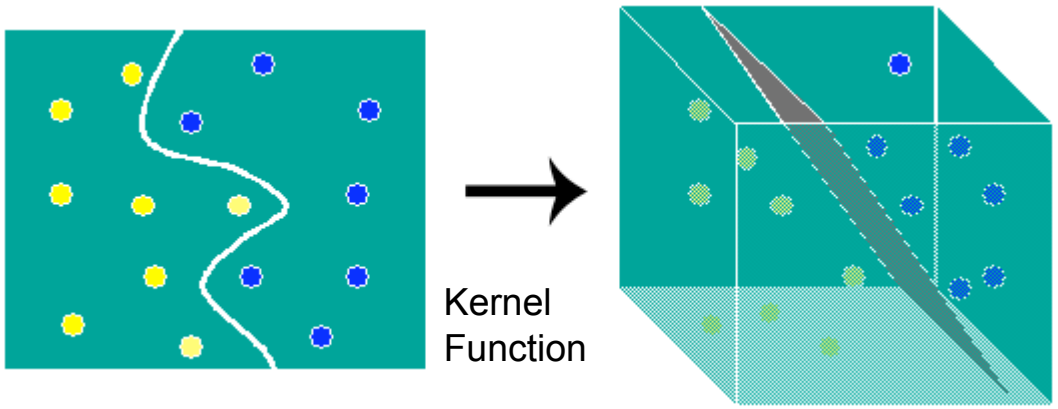


Monitoring System



- Short-term/long-term monitoring
- Count-related/chronological windows
- Different classification and anomaly detection techniques
- Learning-updating/ testing
- Defense against manipulation attacks (poisoning)
- Feature selection and extraction
- Event correlation
- Prevention

Why SVM ?

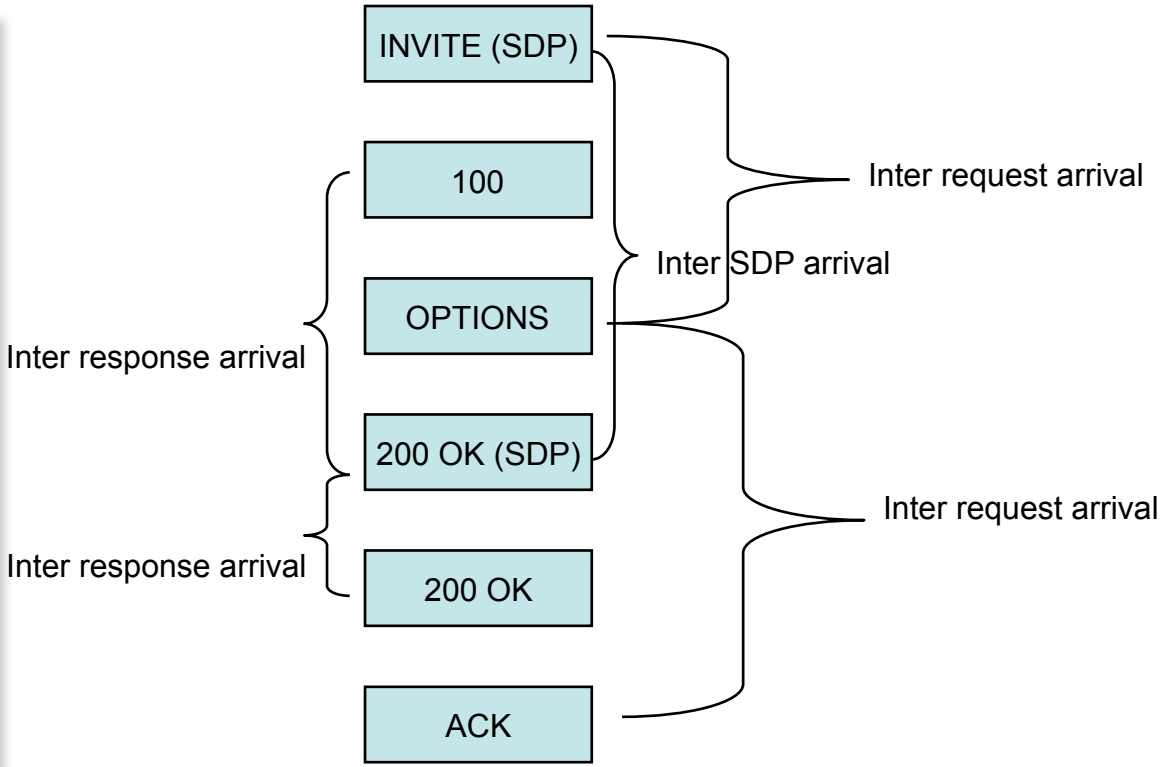


(Radial Basis, Linear, polynomial, sigmoid)

- Known to process high dimensional data
- Classification, regression and exploration of data
- High performance in many domains (Bioinformatics, pattern recognition) and in network-based intrusion detection as well
- Unsupervised Learning

Feature Selection

- We have 38 Features characterizing the SIP traffic
- Distributed over 5 groups:
  1. General statistics
  2. Call-ID based statistics
  3. Dialog final state distribution
  4. Request distribution
  5. Response distribution
- We take into account inbound and outbound messages
- Other features can be investigated as well
- Features must be characterized by a small extraction complexity
- Our feature extraction tool is written in Java using the Jain SIP parser

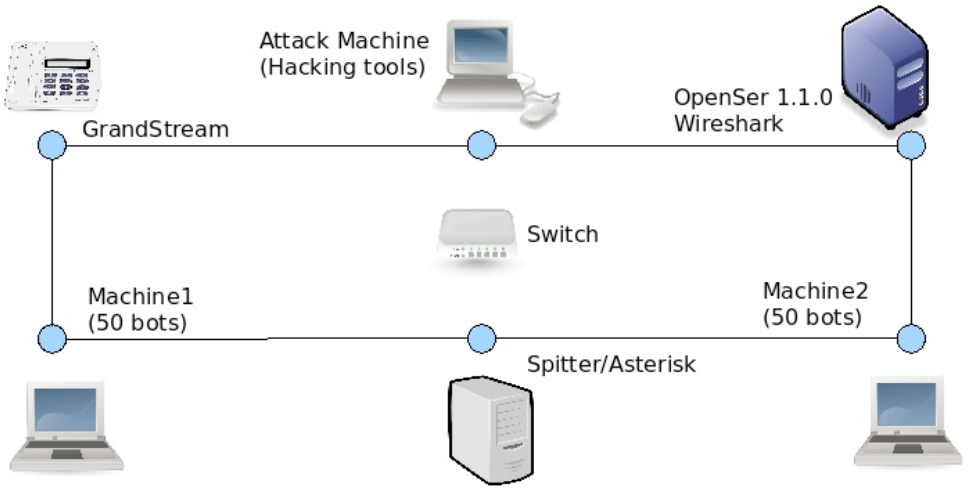


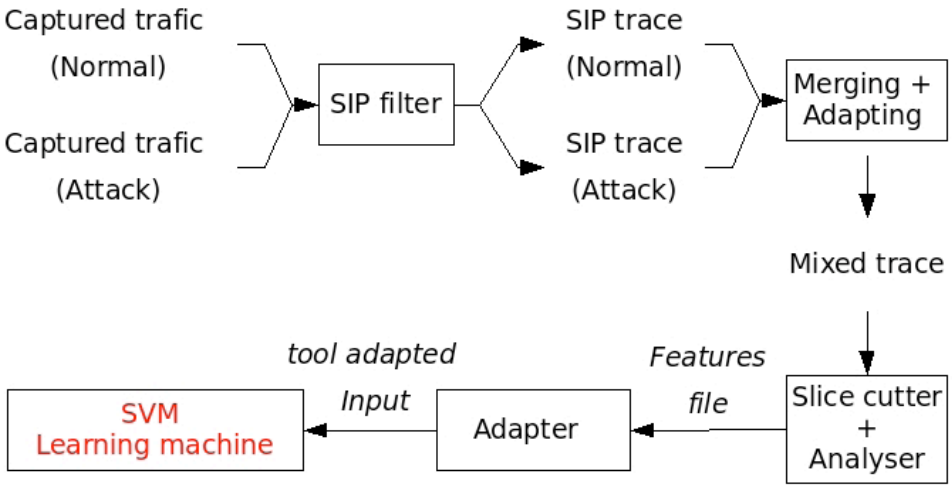
- Average inter request arrival
- Average inter response arrival
- Average inter SDP arrival
- Number of request / total number of messages
- Number of responses /total number of messages
- Number of SDP/ total number of messages
- Number of messages having the same Call-ID

INRIA

Traces and testbed

Real World
VoIP service
provider

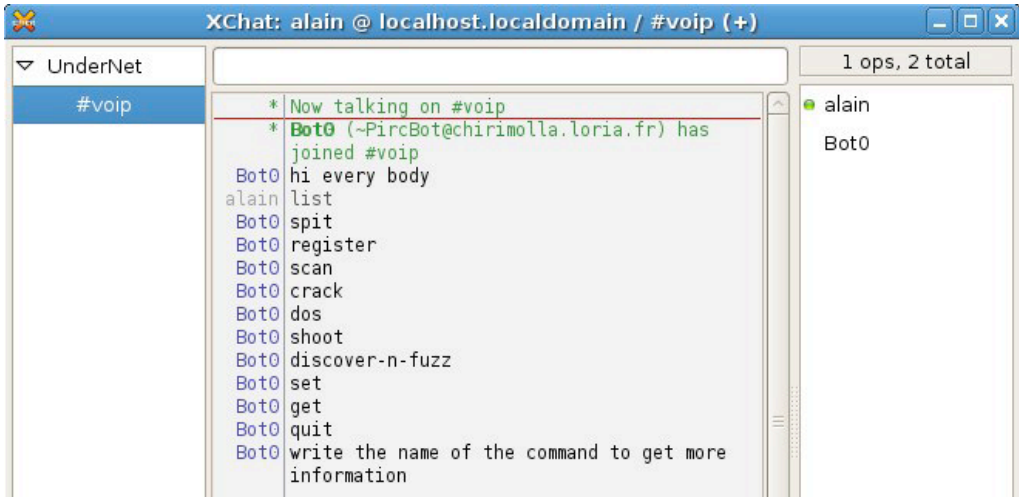


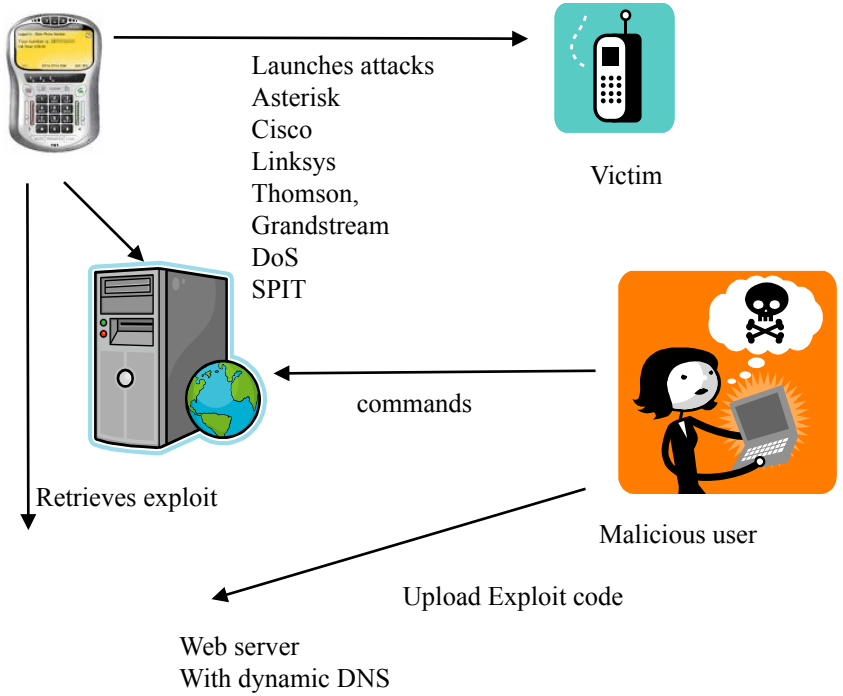


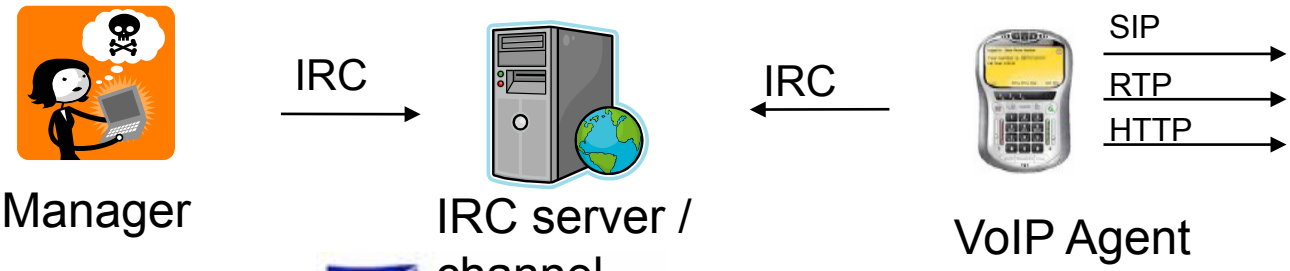
VoIP specific bots

Available from www.loria.fr/~nassar

VoIP Bot



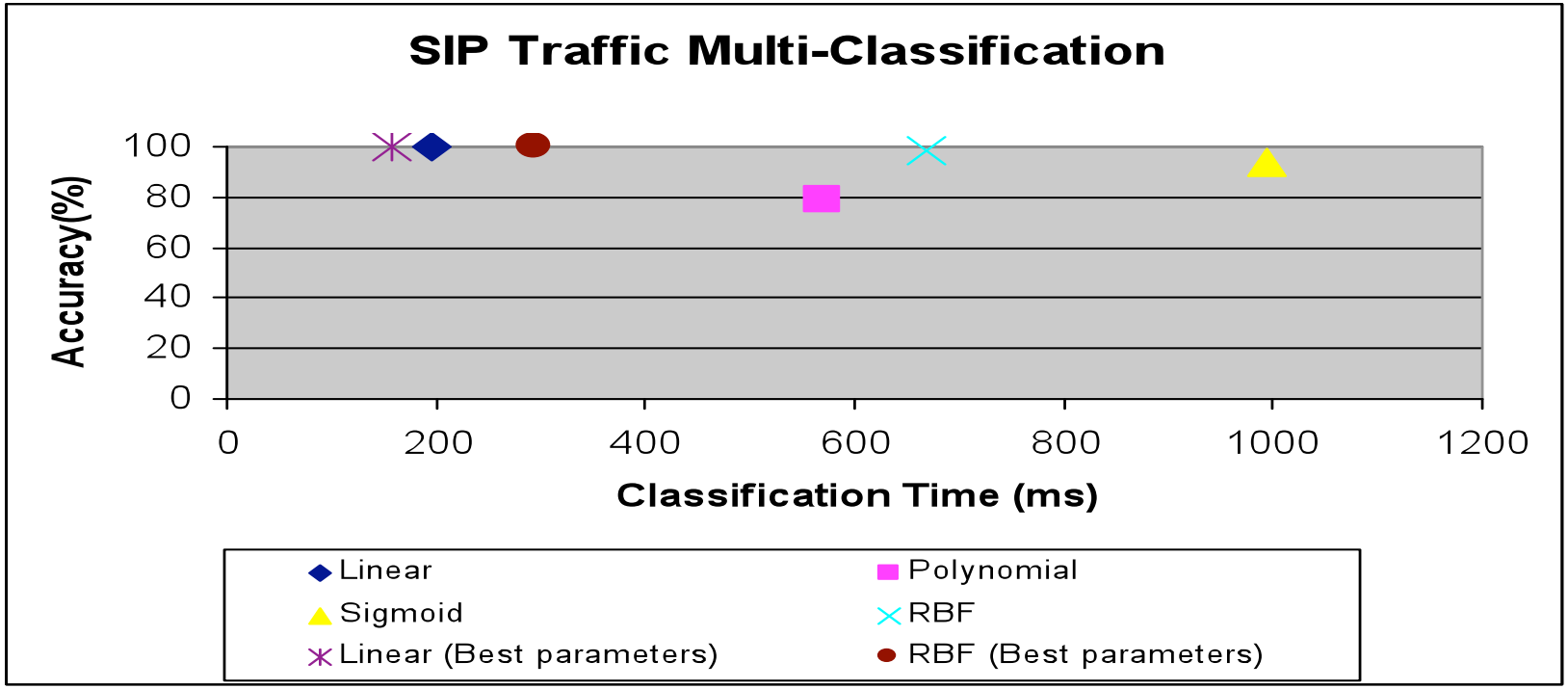




INRIA

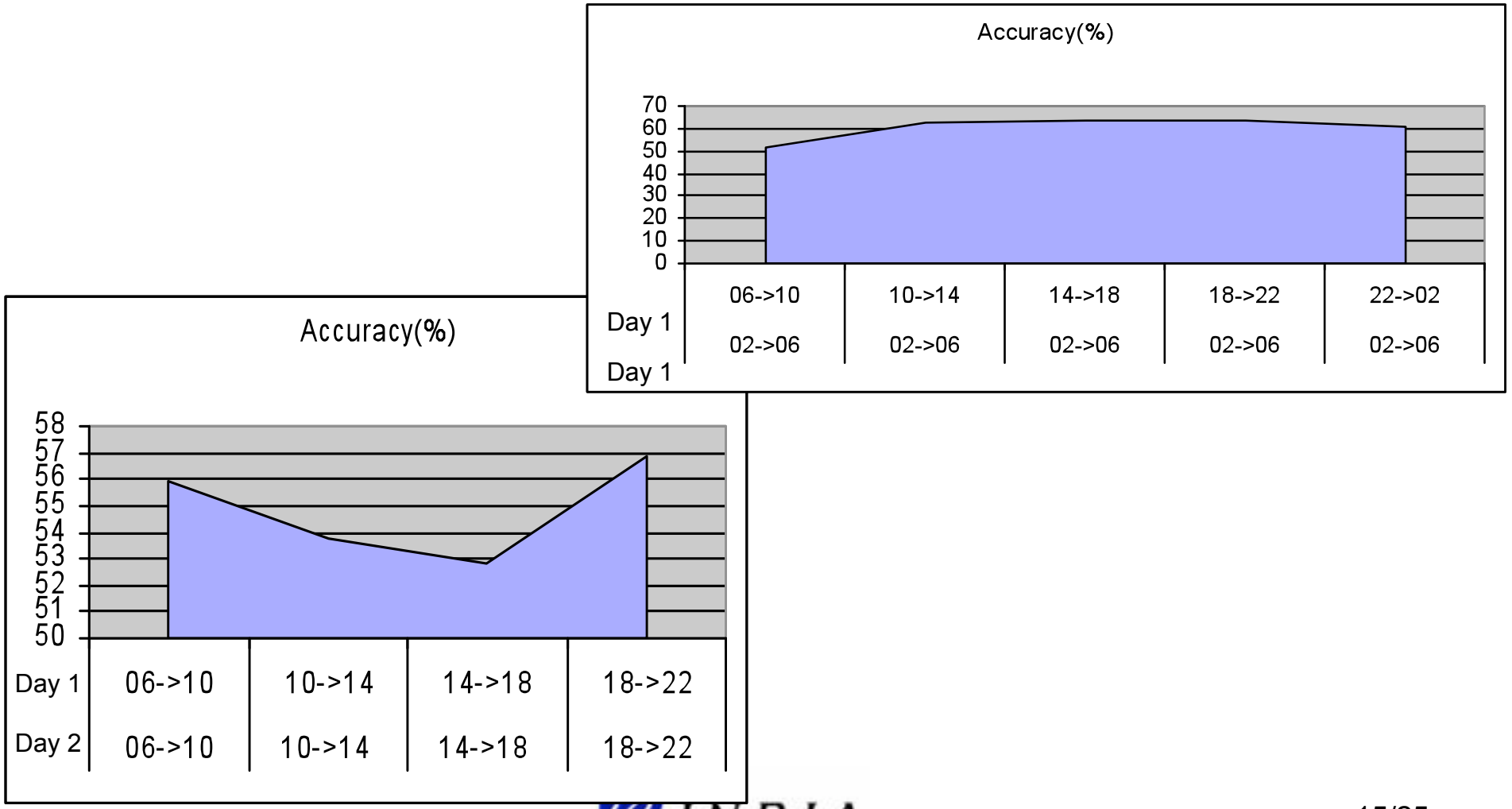
Experiments

Classification time < 1s

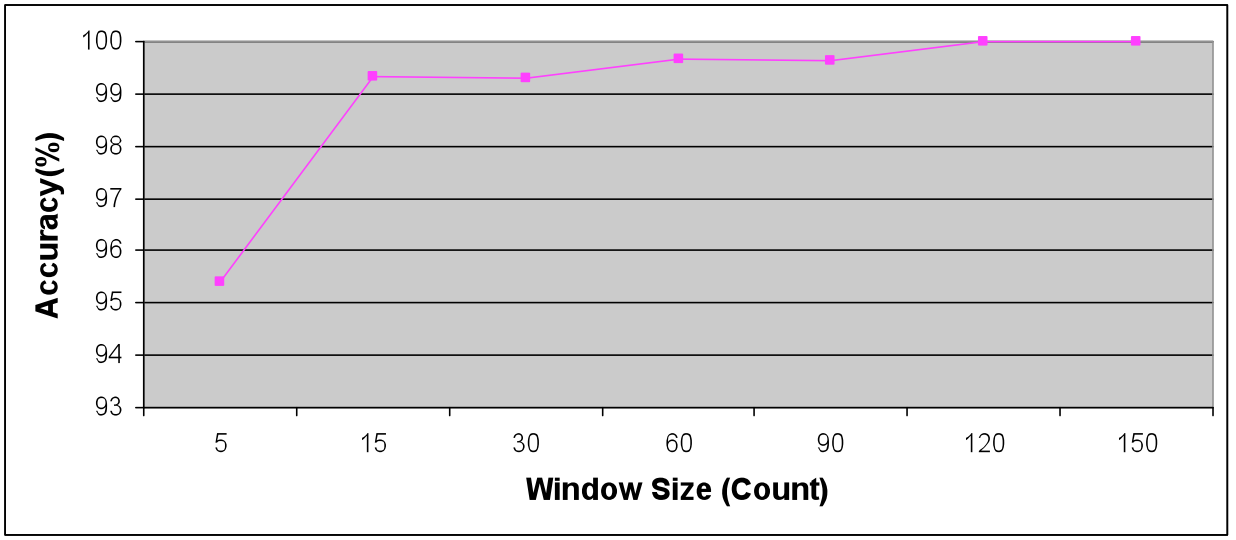


Trace	Normal	DoS	KIF	Unknown
SIP pkts	57960	6076	2305	7033
Duration(min)	8.6	3.1	50.9	83.7

Normal Data Coherence Test

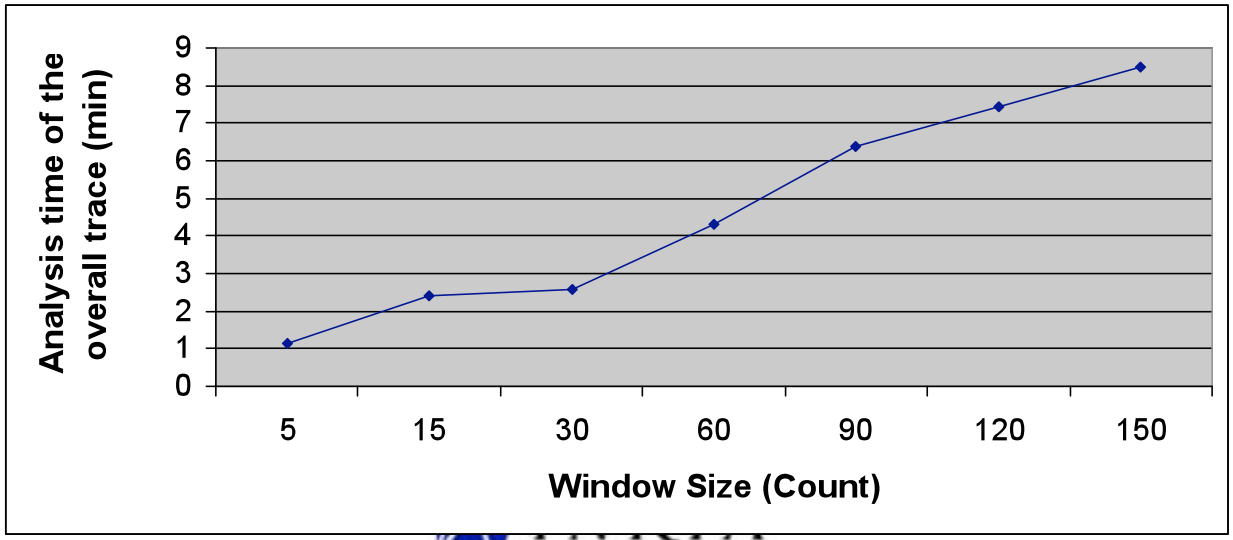


Monitoring Window Size

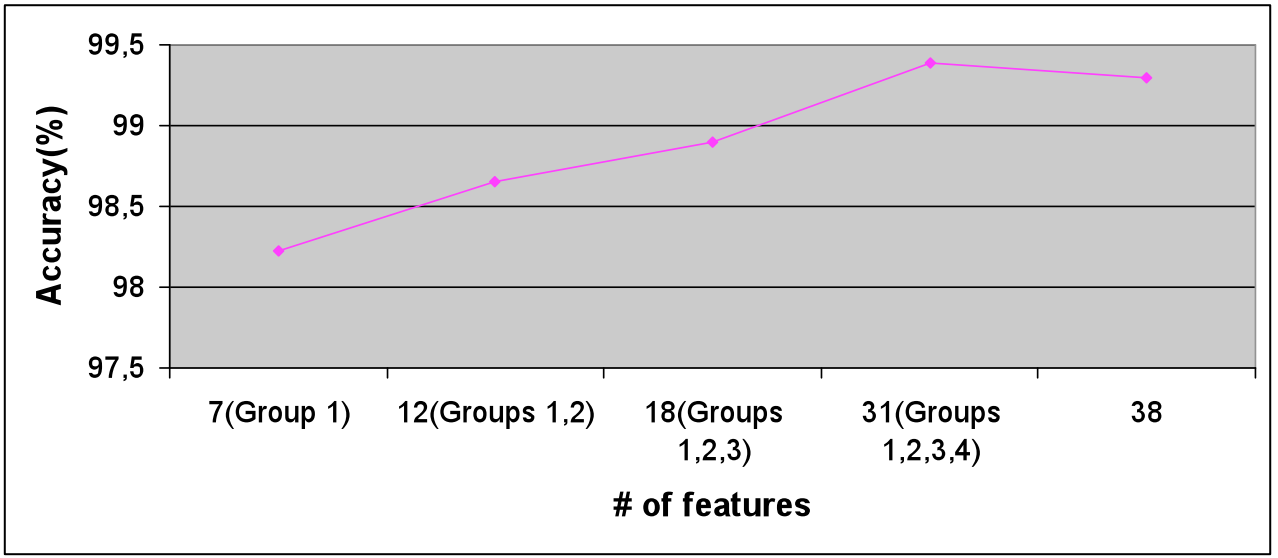


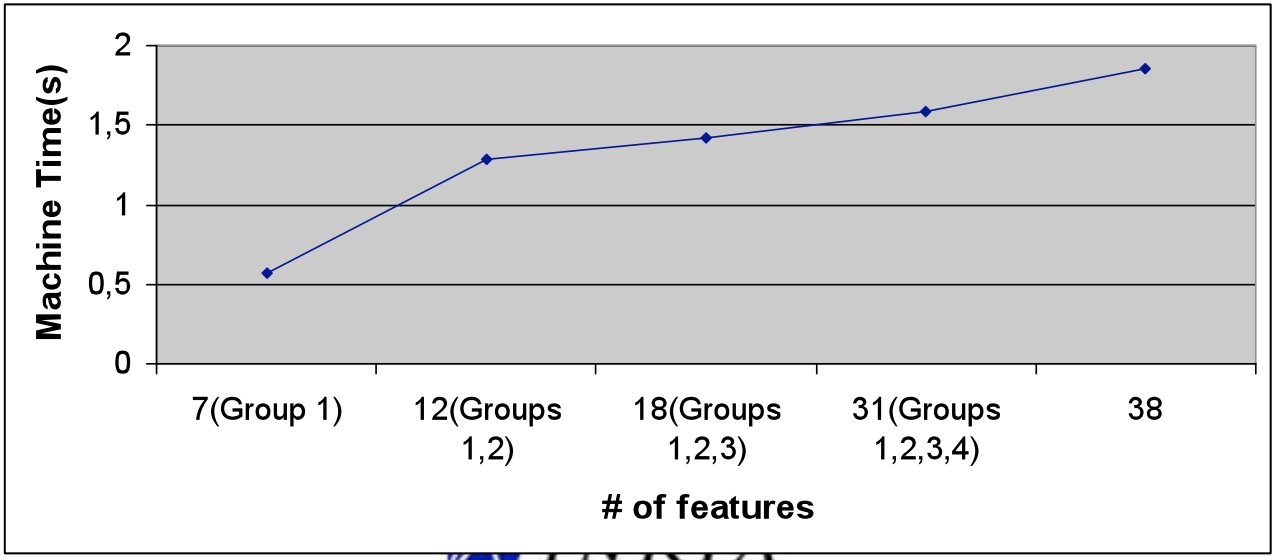
The overall trace is about 8.6 minutes

and message arrival is about 147 Msg/s



Feature selection





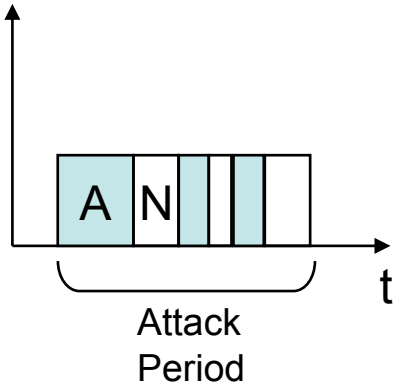
Feature Selection

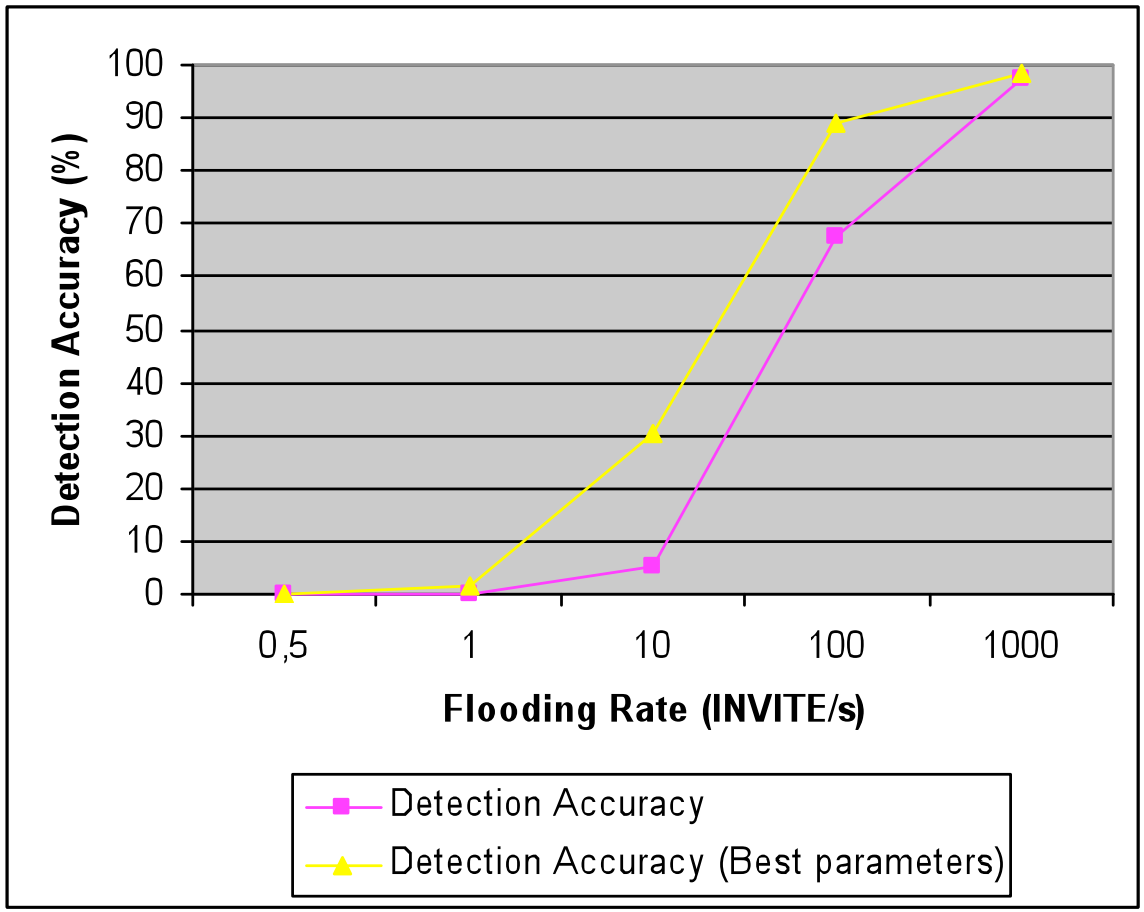
- Greater number of features doesn't mean higher accuracy
- Feature selection increases the accuracy and the performance of the system
- Selected features are highly dependent on the underlying traffic and the attacks to be detected
- A preliminary approach combines F-score and SVM

Flooding Detection

Background traffic ~ 147 Msg/sec

Window = 30 messages





$$\text{Detection\_Accuracy} = \frac{\text{Attack\_period\_detected\_as\_attack}}{\text{Overall\_attack\_period}}$$

INRIA

Selected Features for Flooding / Short Term Monitoring

Number	Name
11	NbReceivers
14	NbCALLSET
20	NbInv
4	NbSdp
2	NbReq
3	NbResp
13	NbNOTACALL
12	AvMsg

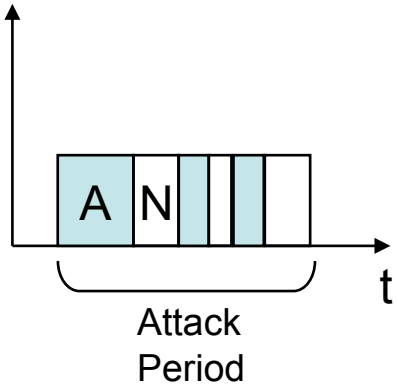
INRIA

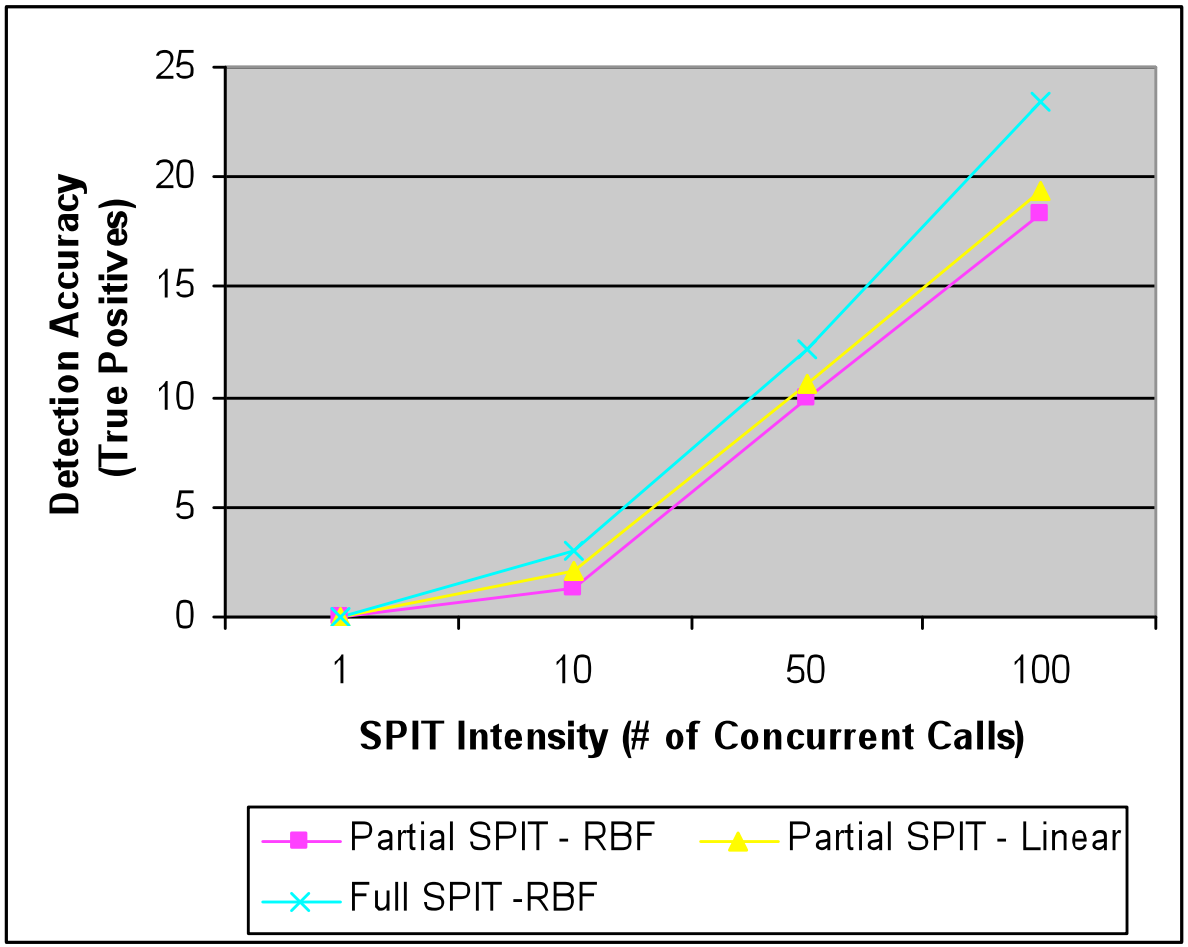
SPIT Detection

Background traffic ~ 147 Msg/sec

Window = 30 messages

False Positive = 0 %





$$\text{Detection\_Accuracy} = \frac{\text{Attack\_period\_detected\_as\_attack}}{\text{Overall\_attack\_period}}$$

INRIA

Selected Features for SPIT / Long Term Monitoring

Number	Name
16	NbRejected
4	NbSdp
20	NbInv
23	NbAck
36	Nb4xx
34	Nb2xx
7	AvInterSdp
35	Nb3xx
13	NbNOTACALL

INRIA

Event Correlation

Predicate	SPIT Intensity
10 Distributed positives in a 2 minutes period	Low (Stealthy)
Multiple Series of 5 successive Positives	Medium
Multiple Series of 10 successive Positives	High

Conclusion and Future works

- Online monitoring methodology is proposed based on SVM learning machine
- Offline experiments shows real-time performance and high detection accuracy
- Anomaly detection and unsupervised learning approach are future works
- Studying traces of other VoIP attacks
- More investigation about the set of features and the selection algorithms
- Extending the event correlation framework in order to reveal attack strategies and attacker plan recognition

INRIA

Annex

Features

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

Features

Group2 - 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

Features

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 COMPLETE/ Total # of Call-ID
19	NbRESIDUE	# of RESIDUE/ Total # of Call-ID

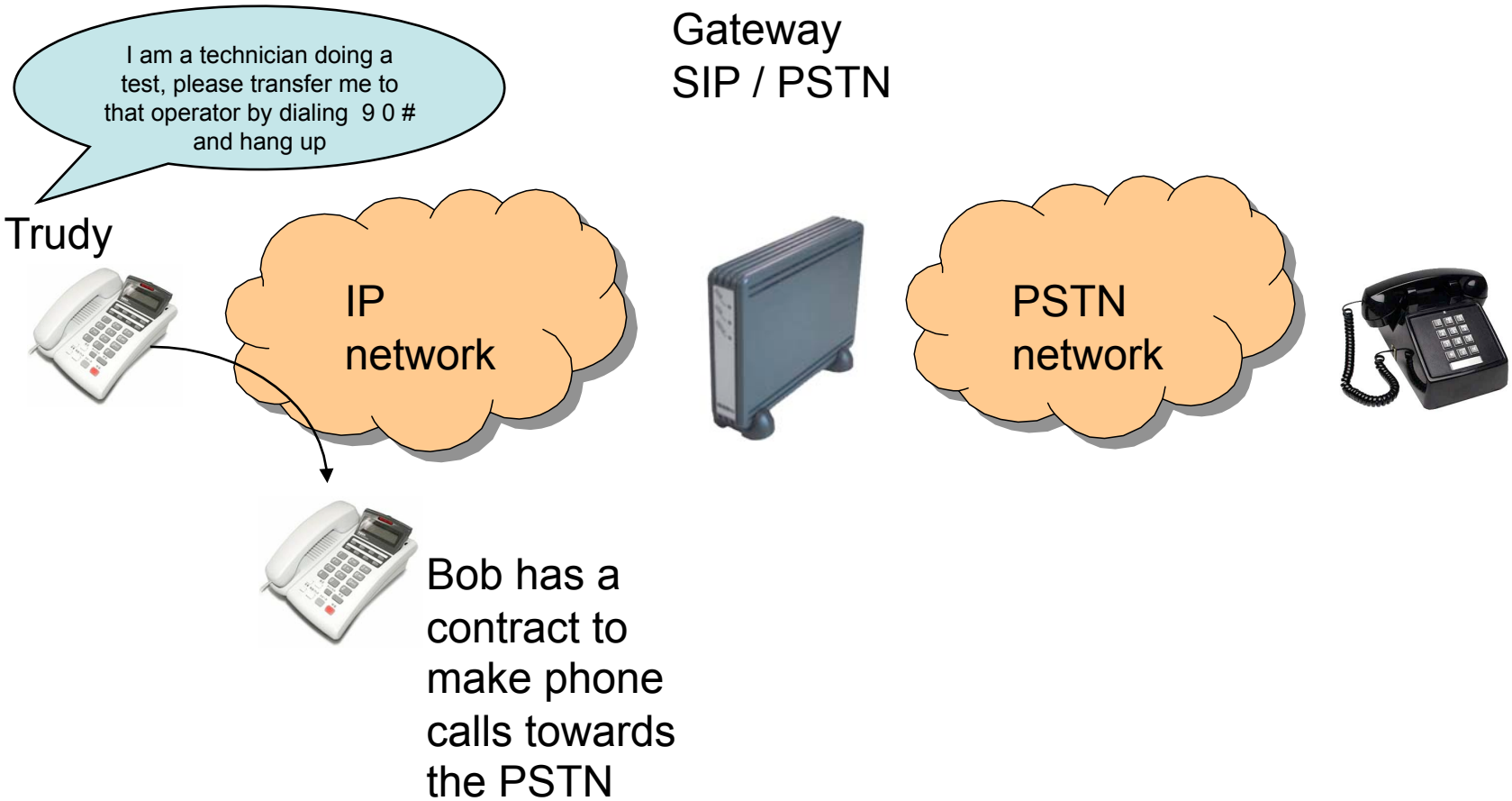
Features

Group 4 – Request 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	NbRef	# 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

Features

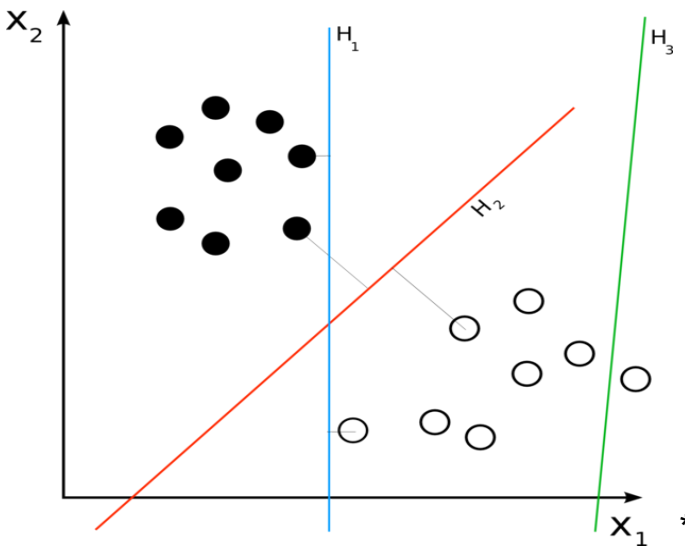
Group5 – Response Distribution		
33	Nb1xx	# of Informational responses / Total # of responses
34	Nb2xx	# of Success responses / Total # of responses
35	Nb3xx	# of Redirection responses / Total # of responses
36	Nb4xx	# of Client error responses / Total # of responses
37	Nb5xx	# of Server error responses / Total # of responses
38	Nb6xx	# of Global error responses / Total # of responses

Phreaking by social engineering scheme

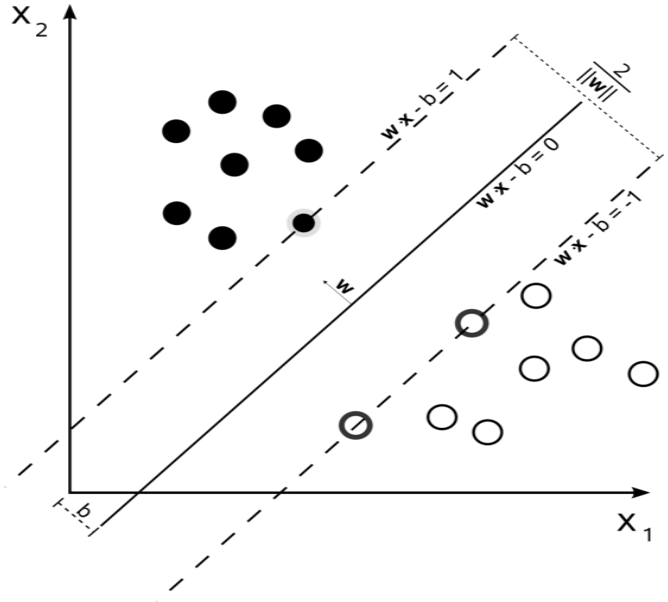


Machine Learning

- Pros
  - Better accuracy, small false alarm rate
  - Compact representation
  - Detecting Novelty
- Cons
  - Embedding of network data in metric spaces
  - Difficulty of getting labels
  - Vulnerable to malicious noise
  - Huge data volumes



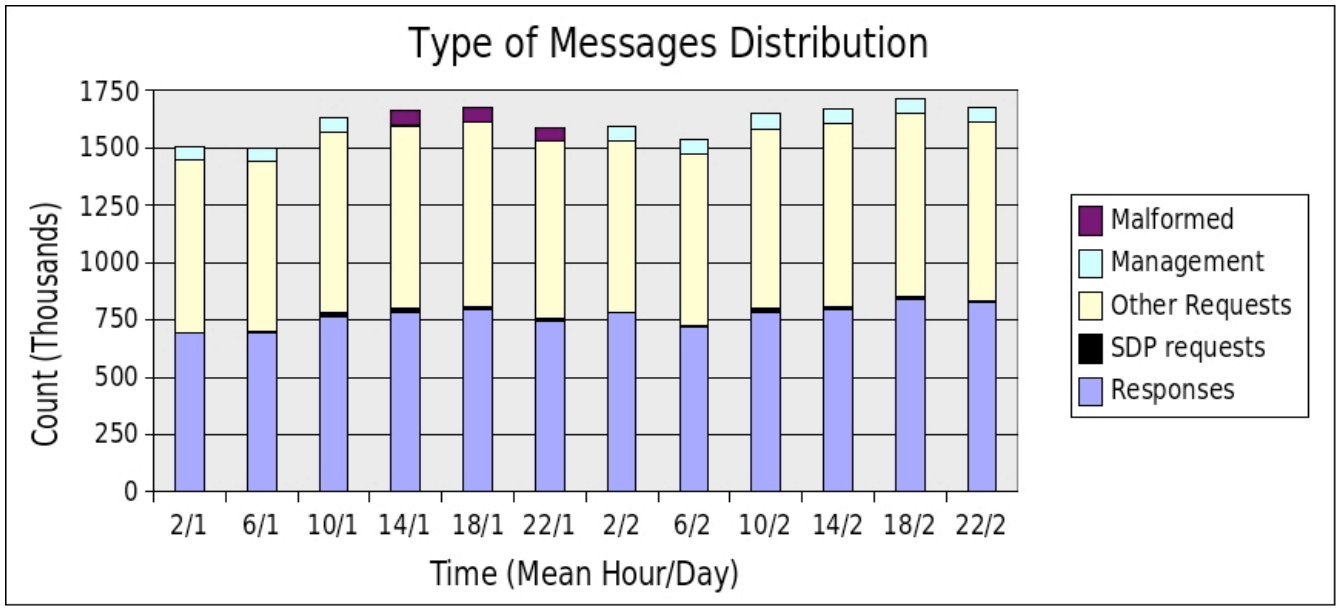
*From Wikipedia



- Linear : $K_l(\vec{x},\vec{z}) = \vec{x}.\vec{z}$
- Polynomial : $K_d(\vec{x},\vec{z}) = (\gamma\vec{x}.\vec{z} + r)^d, \gamma > 0$
- Radial Basis Function : $K_{rbf}(\vec{x},\vec{z}) = \exp(-\gamma|\vec{x} - \vec{z}|^2), \gamma > 0$
- Sigmoid : $K_s(\vec{x},\vec{z}) = \tanh(\gamma\vec{x}.\vec{z} + r), \gamma > 0, r < 0$

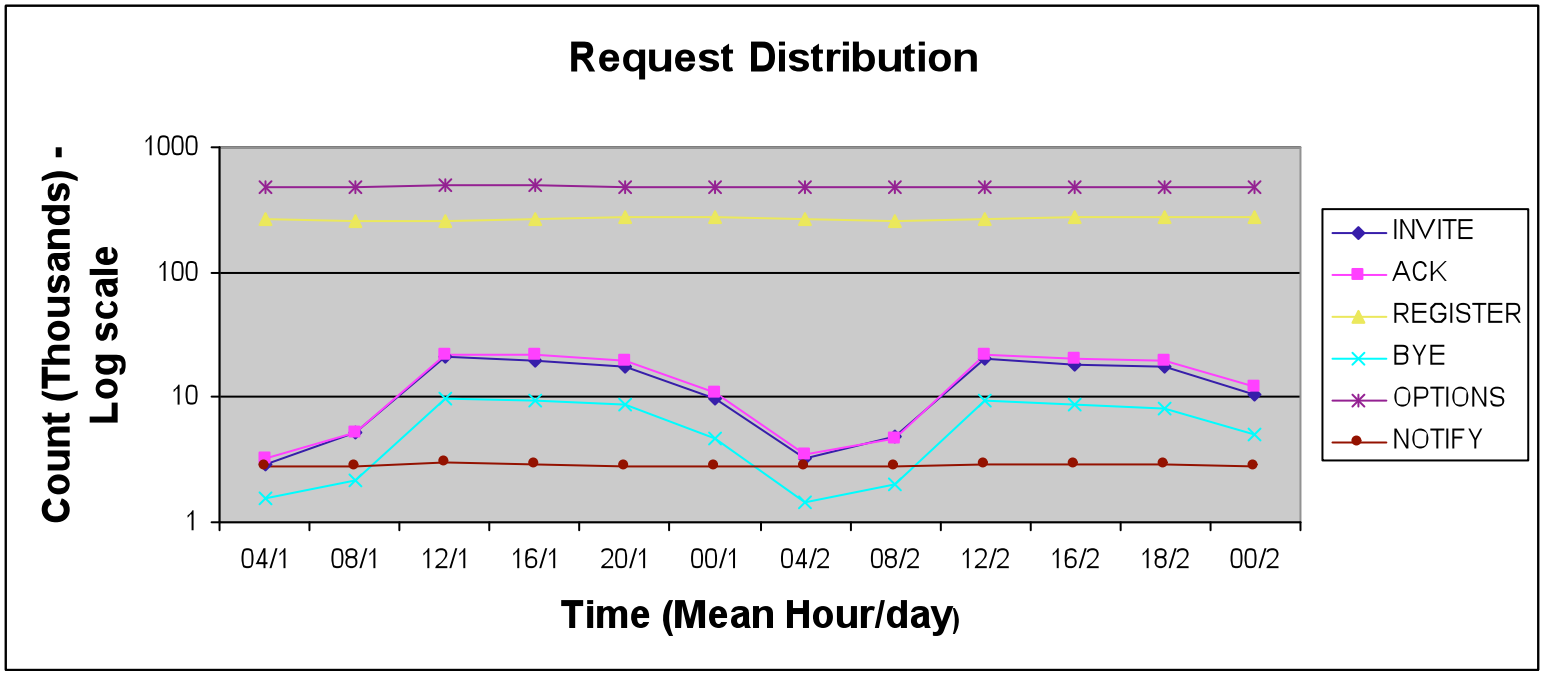
INRIA

Traces



- Call Setup is a small fraction of the signaling traffic
- Some empty messages are used as Ping or KeepALive for device management
- Some messages throw parsing exceptions

Traces

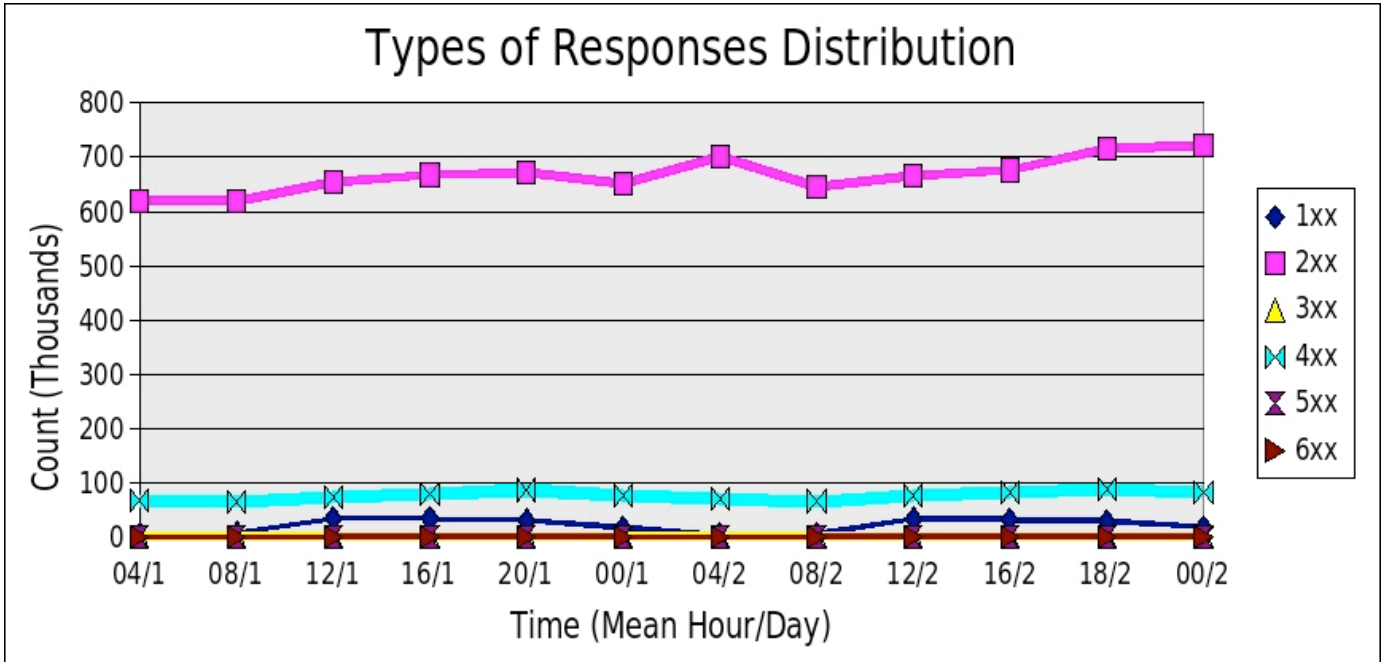


- OPTIONS and REGISTER messages are the most numerous
- MESSAGE, PRACK and UPDATE are absent
- The number of NOTIFY is constant over the time (messages automatically generated at fixed rate)
- #INVITE/#BYE = 2.15 (Not every INVITE result s in a BYE e.g. callee is busy, retransmission, re-INVITE)

•#INVITE/#ACK = 0.92 (Some INVITE are acknowledged twice)

INRIA

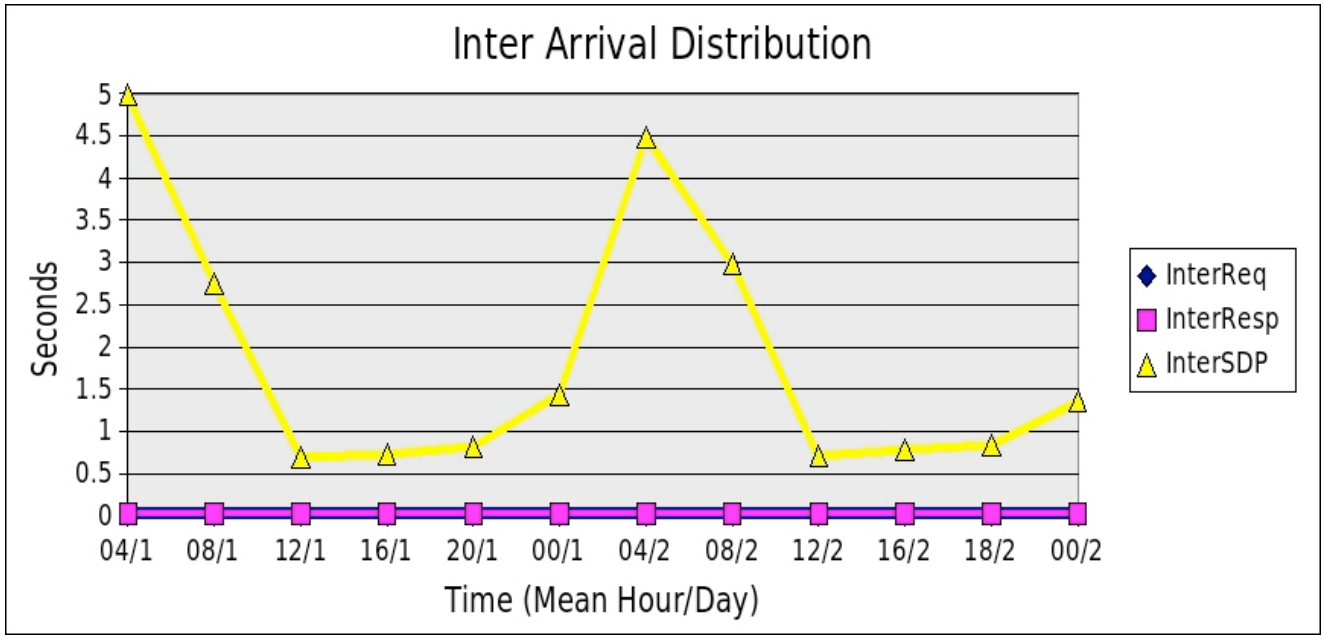
Traces



- The most numerous is the 2xx family (in response to REGISTER and OPTIONS messages)
- #INVITE/#1xx = 0.59 (Probably a 100 Trying and 180 Ringing for each INVITE)

INRIA

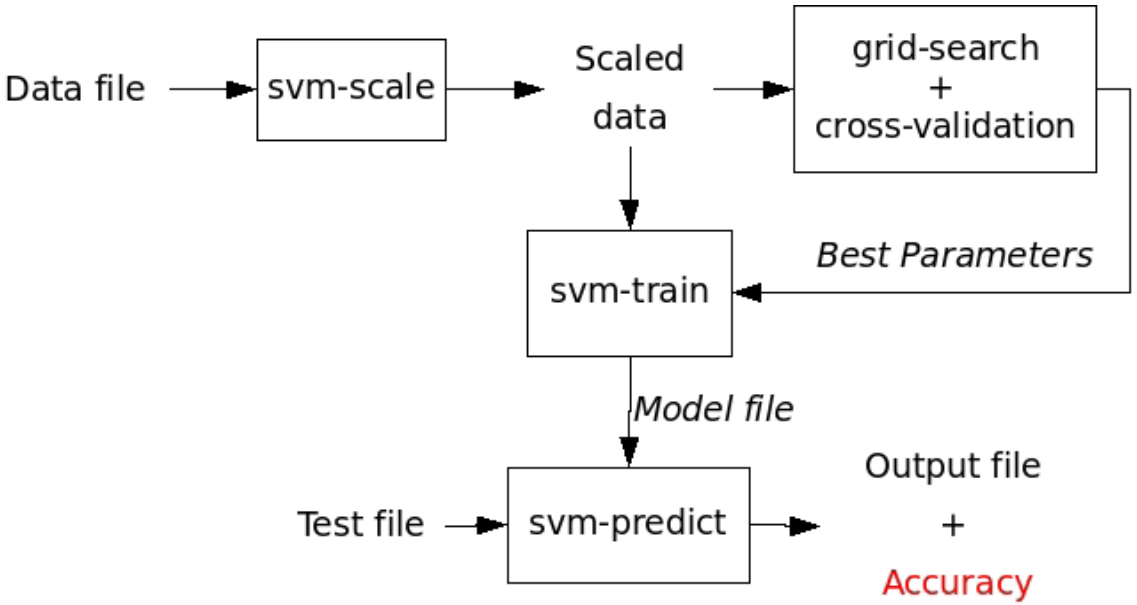
Traces

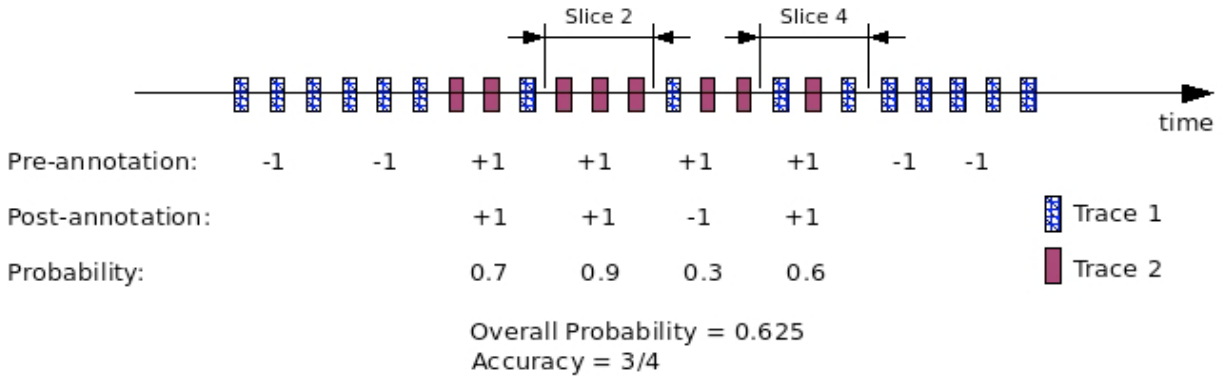


- Average Inter-request = Average Inter Response = 20 ms
- Average inter-request with SDP bodies is inversely proportional to the #INVITE, BYE, ACK and 1xx (which are only used in call-setup)
- Average inter-request carrying SDP reaches 3s in quiet hours and 0.5s in rush hours which reveals a high call-setup traffic

INRIA

LibSVM





INRIA