

## **CASTAFIOR**

**Détection Automatique de Tunnels Illégitimes par  
Analyse Statistique**

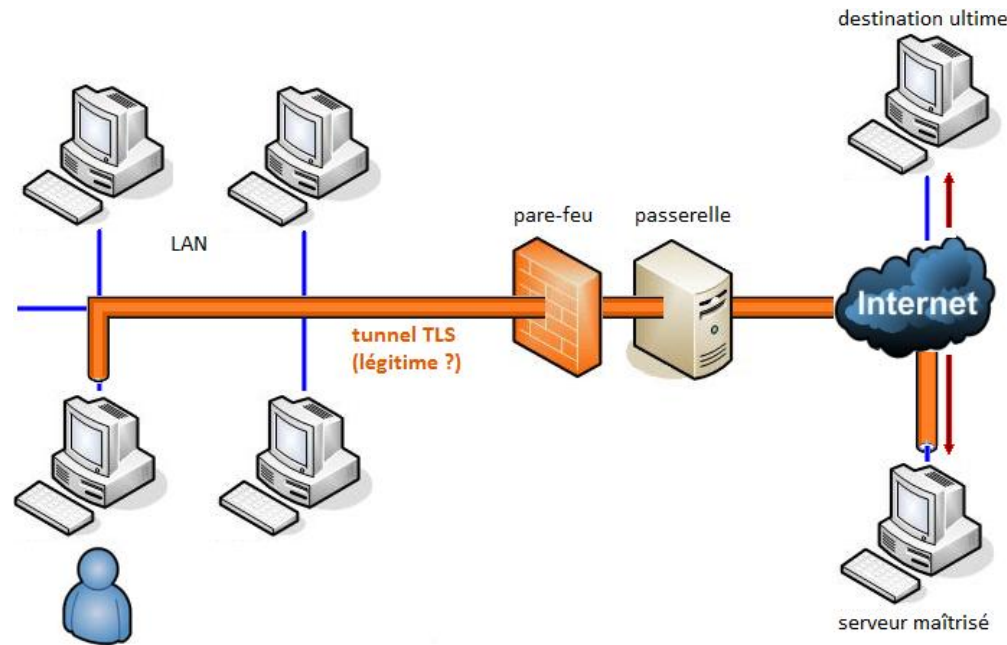
**Fabien ALLARD – Mathieu MOREL**



- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan



- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan



## ■ Le problème

- Contournement possible des pare-feux et des proxys par **tunnel chiffré** (HTTPS) avec un serveur contrôlé
- Les tunnels HTTPS échappent au filtrage par port et à l'inspection des paquets
- Une connexion HTTPS est donc un **canal caché** potentiel, **contournant la politique de sécurité** (ex: encapsulation de sessions TELNET, SSH, P2P, etc.)



- Les méthodes de protection possibles
  - **Cryptanalyser** les flux suspects
    - Impossible ou trop coûteux en pratique
  - Positionner la passerelle en **coupure de chiffrement**
    - + permet une inspection approfondie du *payload* ; utilisé par de nombreuses entreprises
    - sensibilisation de la passerelle ; pas de chiffrement de bout-en-bout ; risque d'engorgement
  - Reconnaître le protocole à l'origine du flux par **analyse statistique/comportementale**
    - + Utilisation des paramètres observables après chiffrement pour classer les flux ; rapide ; pas de déchiffrement ; respect des données privées
    - Peut générer de fausses alertes ; possibilités de contournement par mimétisme de flux autorisés
  
- L'objectif
  - Diminution du risque d'un « canal caché » mais pas suppression ;
  - **Preuve de concept** de la technologie : analyse de la faisabilité et des limites.



- La technologie pourrait permettre de **réduire les débits des canaux cachés** par encapsulation ;
  - Détection non fiable à 100%
  - Méthodes d'évitement possibles, mais contraignantes vis-à-vis du débit utile dans l'encapsulation
- Permet la **détection de logiciels P2P** « furtifs » ou de **backdoors** communiquant avec l'extérieur par tunnel HTTPS ;
- Cadre d'emploi : sur les passerelles
  - côté réseau local d'un réseau d'entreprise ;
  - ou en amont des chiffreurs IP d'un réseau de défense ;
- Constitue un complément intéressant aux **IDPS** actuels (Arkoon, NetASQ, etc.) ;
- Peut être intégré dans une architecture de sécurité plus vaste, en complément d'information.



- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan



- Objectif : déterminer le protocole à l'origine d'un flux chiffré
- Hypothèse : chaque protocole (*TELNET, SSH, P2P, videolan, HTTP, etc.*) a un **comportement caractéristique**
  - Echanges de paquets
  - Tailles de paquets
  - Temps inter-paquets
- Cette **empreinte** reste observable après chiffrement/encapsulation dans HTTPS
- Idée :
  - Extraire les **paramètres pertinents** des flux chiffrés
  - Comparer ces paramètres à un **modèle statistique** préalablement construit identifier le **protocole le plus probable**
- Contournement du système par **mimétisme d'un flux légitime**
  - Impose de **fortes contraintes** à l'attaquant





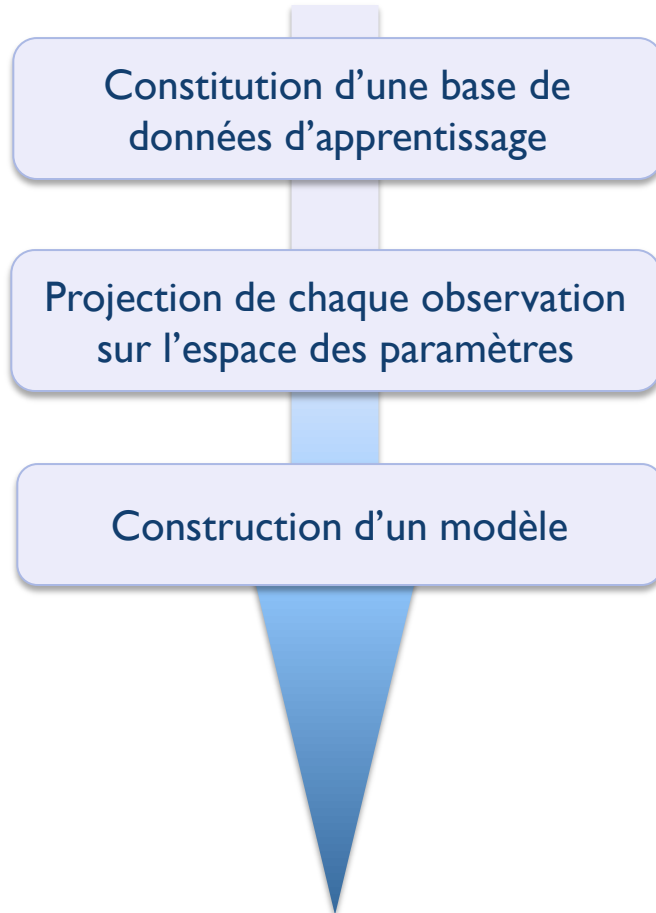
Constitution d'une base de données d'apprentissage



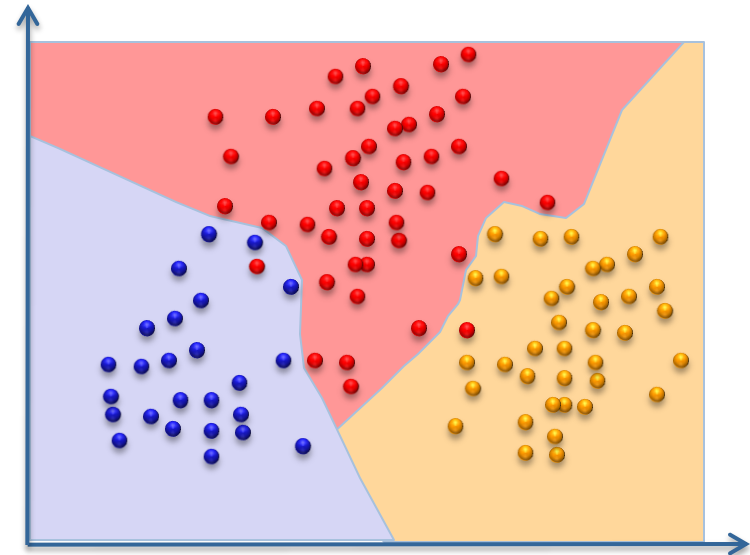
No.	Time	Source	Destination	Protocol	Info
5452	610.72074	192.168.1.102	192.168.1.1	TCP	2459 > http [ACK] Seq=20 Ack=...
5453	610.72121	192.168.1.102	192.168.1.1	TCP	2459 > http [FIN, ACK] Seq=20 Ack=...
5454	610.72195	192.168.1.1	192.168.1.102	TCP	http > 2459 [RST, ACK] Seq=5454 Win=0 Len=0
5455	610.76519	192.168.1.102	192.168.1.4	DNS	Standard query A www.mikrotik.com
5456	610.76856	192.168.1.4	192.168.1.102	DNS	Standard query response A 66
5457	611.09368	192.168.1.102	192.168.1.3	NBNS	Name query NBSTAT *<00><00><00>
5458	611.09386	192.168.1.3	192.168.1.102	NBNS	Name query response NBSTAT
5459	611.46980	192.168.1.102	192.168.1.4	NBNS	Name query NBSTAT *<00><00><00>
5460	611.47022	192.168.1.4	192.168.1.102	NBNS	Name query response NBSTAT
5461	612.88506	Cisco-Li_a9:b5:eb	Broadcast	ARP	who has 192.168.1.200? Tell me
5462	613.02826	192.168.1.102	192.168.1.245	SNMP	get-next-request
5463	613.03054	192.168.1.245	192.168.1.102	SNMP	get-response
5464	613.73557	192.168.1.4	192.168.1.102	synerg	24800 > 1168 [PSH, ACK] Seq=1168 Win=0 Len=0
5465	613.73678	192.168.1.102	192.168.1.4	synerg	1168 > 24800 [PSH, ACK] Seq=1168 Win=0 Len=0
5466	613.73752	192.168.1.4	192.168.1.102	synerg	24800 > 1168 [PSH, ACK] Seq=1168 Win=0 Len=0
5467	613.73793	192.168.1.102	192.168.1.4	synerg	1168 > 24800 [PSH, ACK] Seq=1168 Win=0 Len=0

Frame 1 (46 bytes on wire, 46 bytes captured)  
Ethernet II, Src: AsustekC\_24:50:9e (00:0e:a6:24:50:9e), Dst: LinksysG\_d8:68:b5 (00:0c:4d:00:00:00)  
Internet Protocol, Src: 192.168.1.102 (192.168.1.102), Dst: 192.168.1.245 (192.168.1.245)

```
0000 00 0c 41 d8 68 b5 00 0e a6 24 50 9e 08 00 45 00 ..A.h...$.P...E.
0010 00 20 6d 9c 00 00 40 01 88 95 c0 a8 01 66 c0 a8 . m...@. ....f..
0020 01 f5 08 00 cf bf 28 0e 4e 39 b8 e6 f9 11 .....(.N9....
```

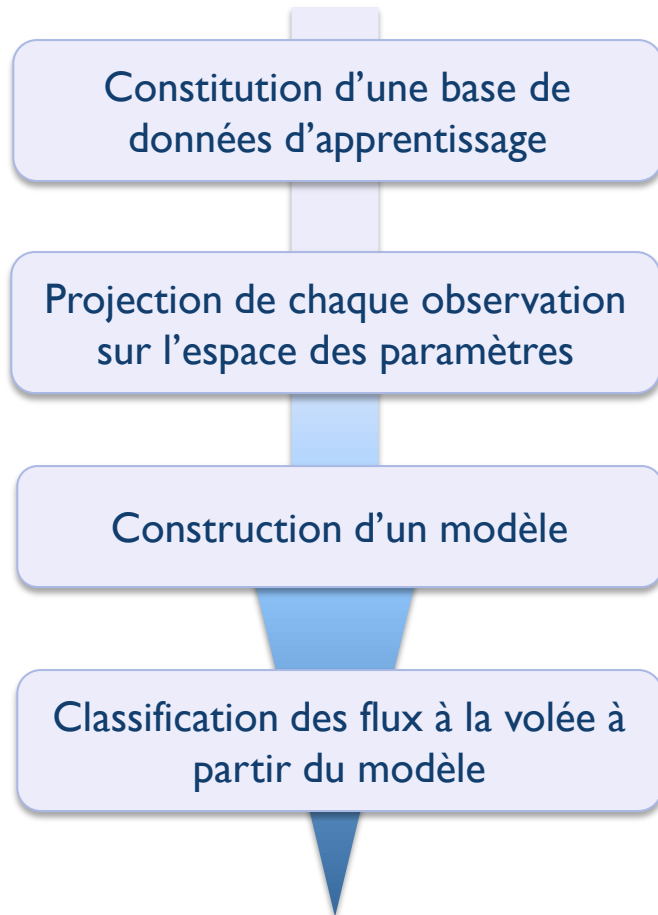


Nombre paquets transmis  
client vers serveur

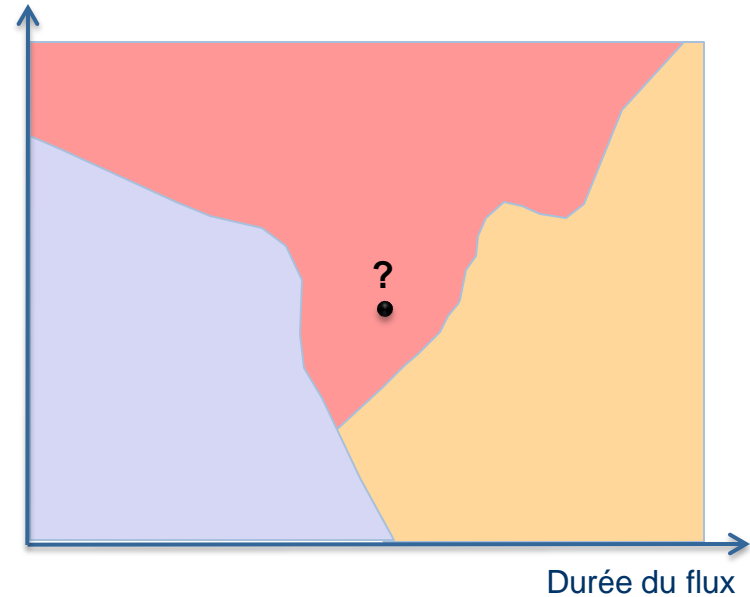


Durée du flux

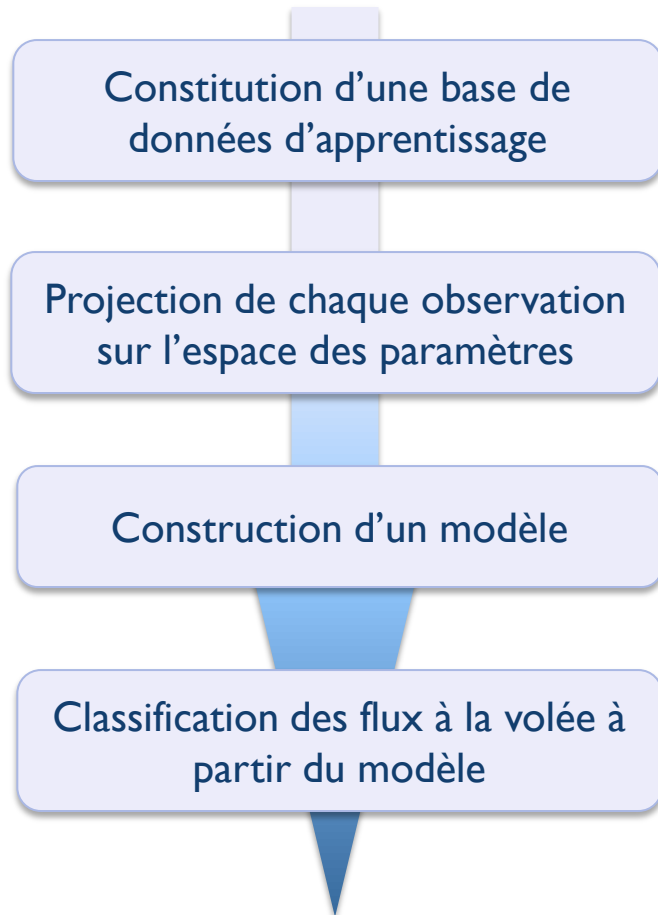
- flux SSH
- flux HTTPS
- flux P2P



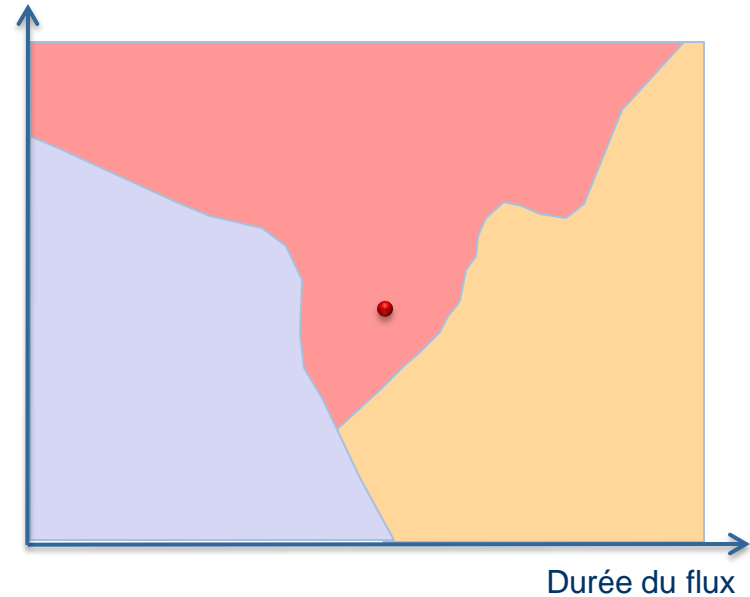
Nombre paquets transmis client vers serveur



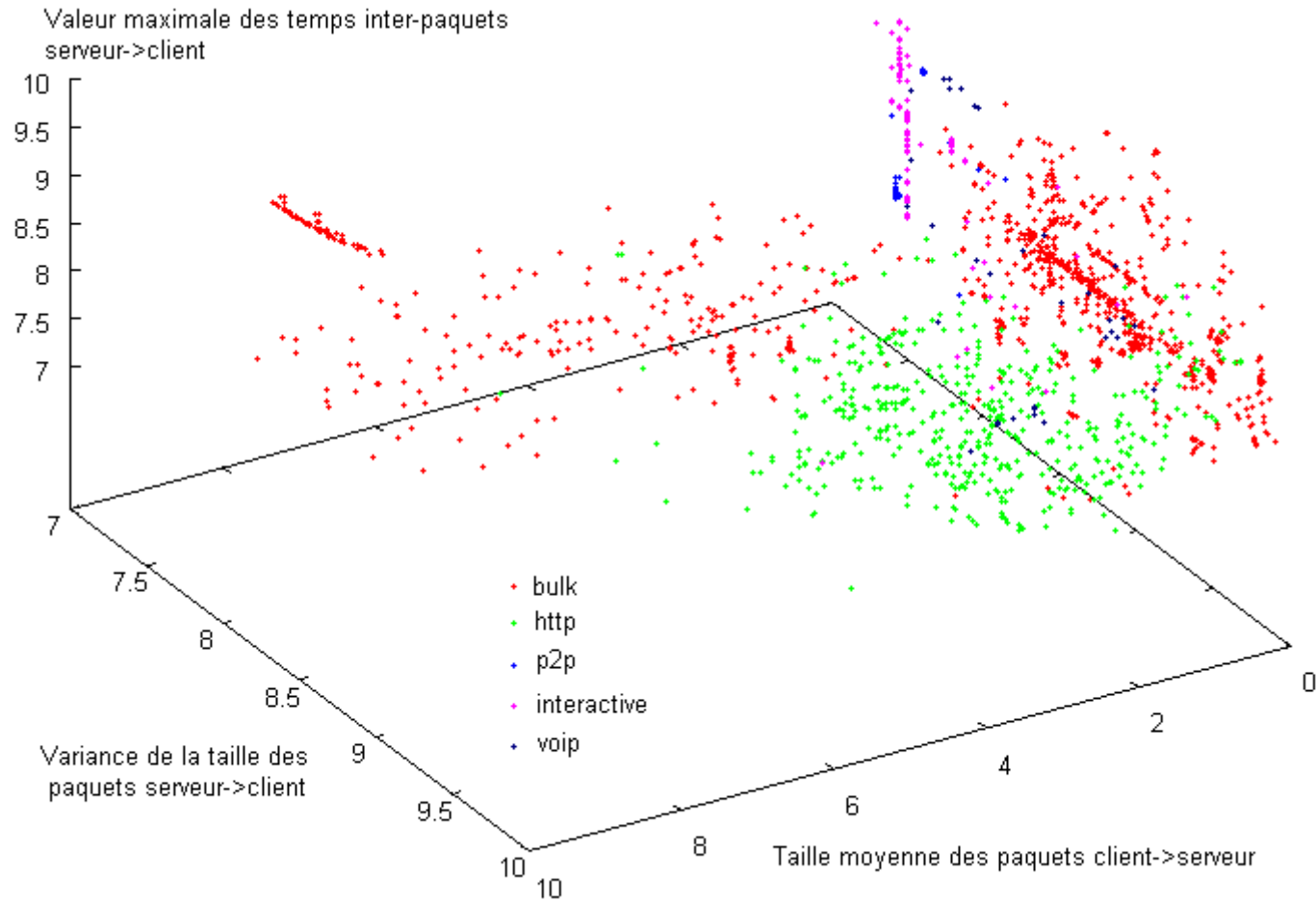
- flux SSH
- flux HTTPS
- flux P2P



Nombre paquets transmis  
client vers serveur



- flux SSH
- flux HTTPS
- flux P2P

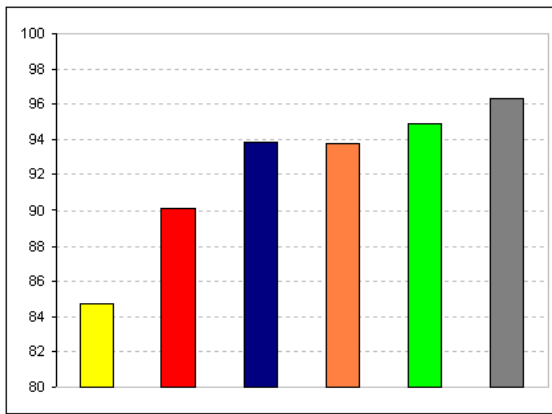


Projection de quelques flux réels sur trois paramètres utilisés pour la classification

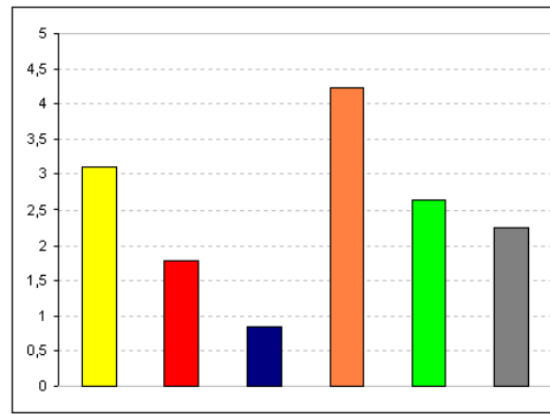
# Choix d'une méthode de classification « macroscopique »



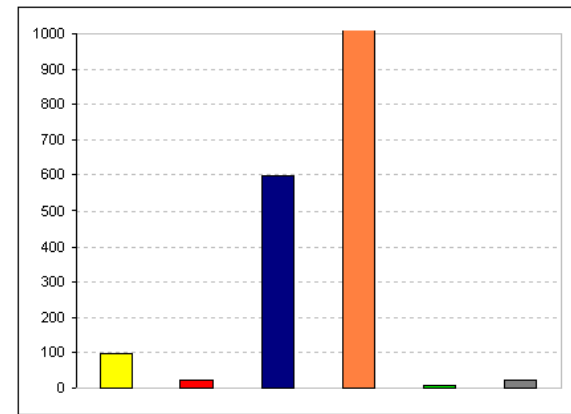
- Utilisation d'une **base de données publique** de 20 000 flux clairs
- Détermination des **10 paramètres à extraire** des flux
  - Algorithmes de sélection automatique des paramètres les plus discriminants et les moins corrélés
- A partir de l'étude bibliographique, applications de 6 méthodes d'apprentissage
  - **Méthode de Bayes naïve**
  - **Support Vector Machine (SVM)**
  - **Arbre de décision C4.5**
  - **Méthode RandomForest**
  - **Méthode K-Means**
  - **Mélange de gaussiennes (GMM)**



Taux de bonne classification

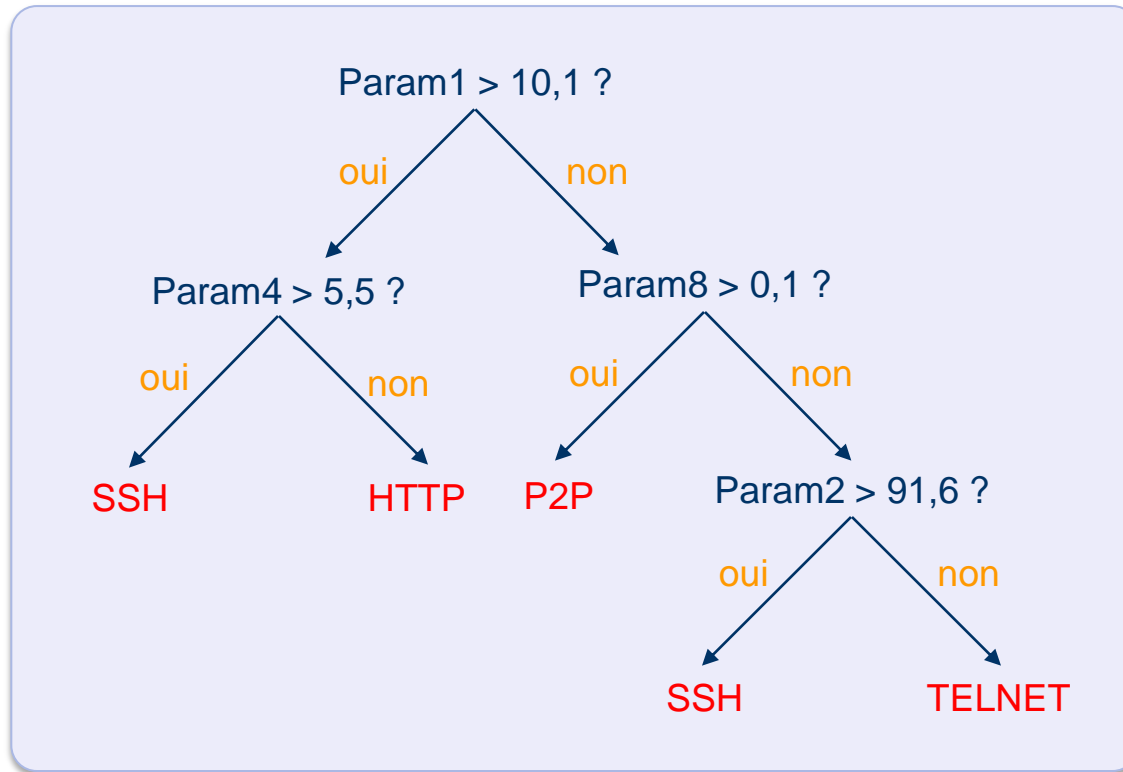


Taux de faux positifs (HTTP mal classés)



Temps de classification (en µs)

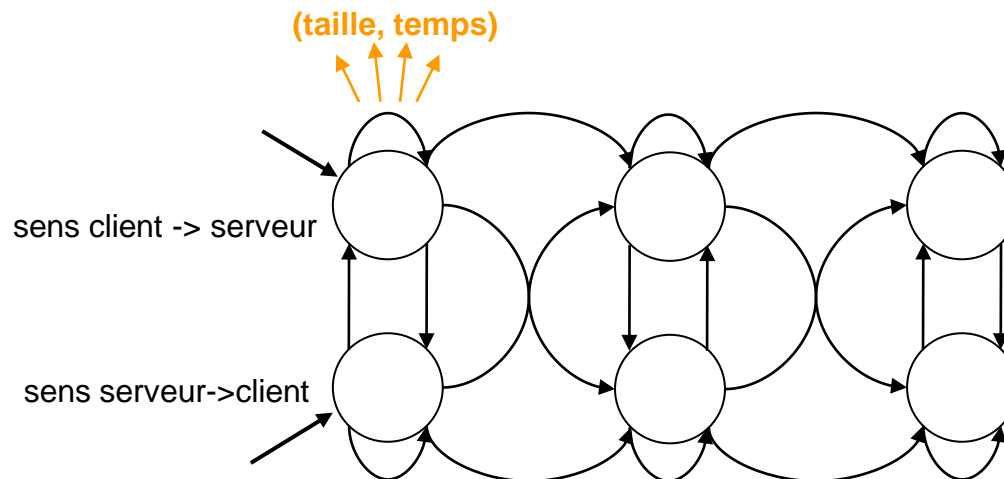
- Qu'est-ce qu'un arbre de décision ?



- *RandomForest* est une forêt d'arbres de décision aléatoires, on construit  $n$  arbres de décision différents, puis le classement s'effectue par vote majoritaire



- Objectif : exploiter l'information séquentielle des enchaînements de paquets
- Outil : les Modèles de Markov Cachés (HMM), proches des automates
- Chaque HMM modélise un protocole.

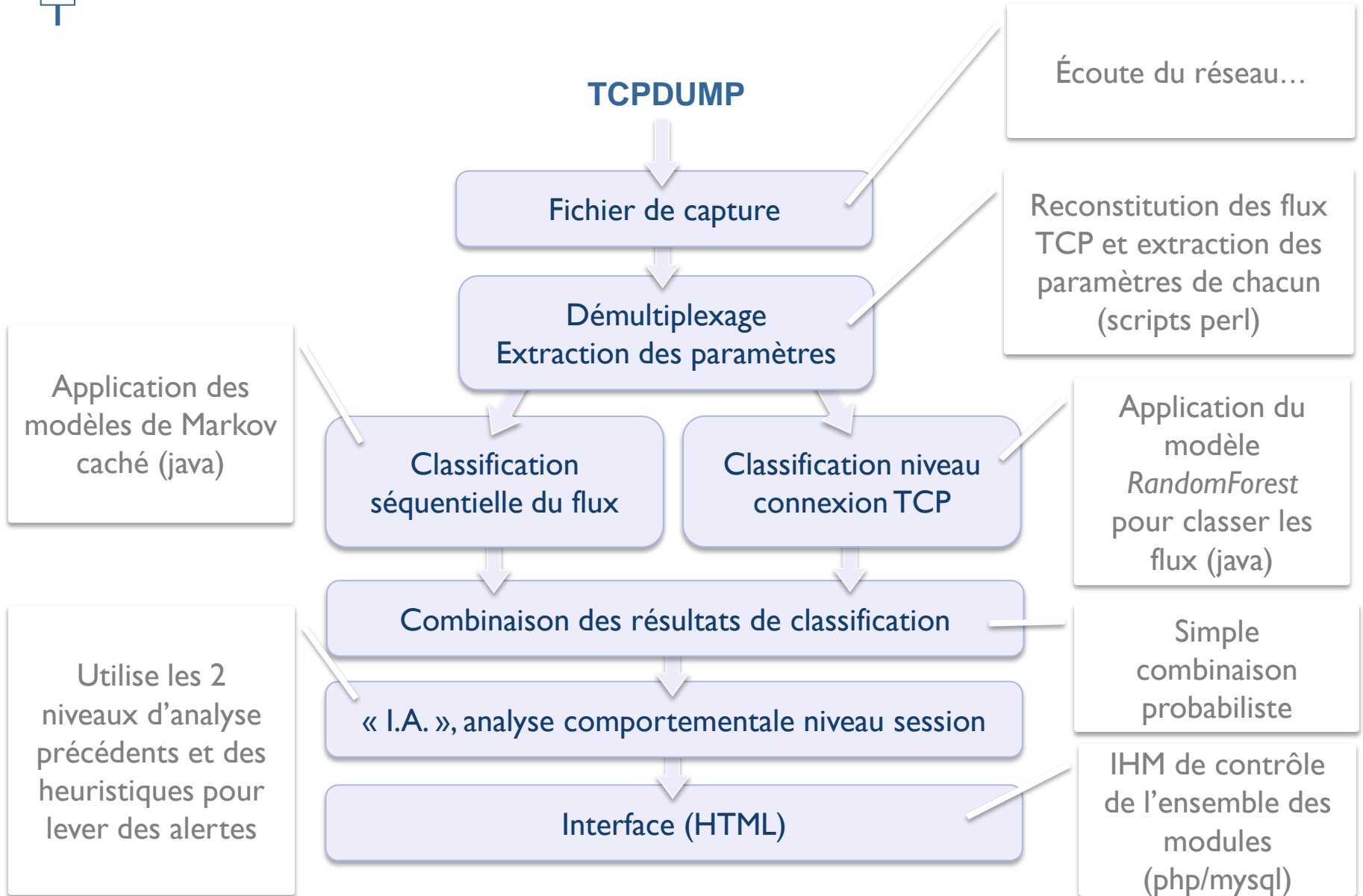


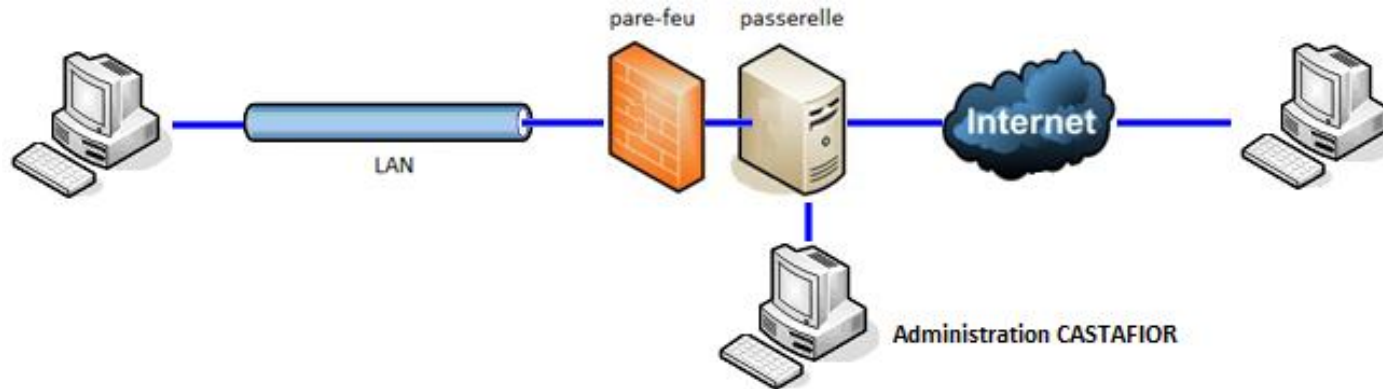
- Apprentissage (à partir de la base d'apprentissage)
  - Calcul des probabilités de transition et d'émission par l'algorithme de Baum Welsh
- Classification
  - On calcule sa probabilité d'émission par chacun des HMM (algorithme de Viterbi)
  - On retient le HMM ayant donné la probabilité maximale





- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan





http://localhost/capture/lancer.php

## Classification automatique de flux par analyse statistique

### Lancer une nouvelle capture...

Cette page permet de commencer une nouvelle écoute du trafic, et de visualiser en temps réel la classification des flux.

Timestamp	IP client	IP serveur	Port client	Port serveur	Classe	Conf.
02/11/09 17:11:39	192.168.0.1	192.168.1.1	60487	443	HTTPs	1
02/11/09 17:11:16	192.168.0.1	192.168.1.1	45886	51025	MAILs	0.95
02/11/09 17:11:59	192.168.0.1	192.168.1.1	49258	51110	MAILs	1
02/11/09 17:11:11	192.168.0.1	192.168.1.1	54515	51022	SSHs	0.75
02/11/09 17:11:08	192.168.0.1	192.168.1.1	54513	51022	SSHs	0.8
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60458	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60457	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60456	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60455	443	HTTPs	1
02/11/09 17:11:22	192.168.0.1	192.168.1.1	35317	51110	MAILs	0.65

#### Sortie console

```
Mode de classification temps reel...
Lancement du demultiplexeur...
Lancement du script d'extraction des parametres...
Lancement du classificateur...
Utilisation du jeu de parametres : 2
Traitement d'un nouveau flux (1)...
Traitement d'un nouveau flux (2)...
Traitement d'un nouveau flux (3)...
Traitement d'un nouveau flux (4)...
Traitement d'un nouveau flux (5)...
Traitement d'un nouveau flux (6)...
Utilisation du jeu de parametres : 2
```

#### Rapport d'analyse

02/11/09 192.168.0.1->192.168.1.1 : detection d'une encapsulation probable...

02/11/09 192.168.0.1->192.168.1.1 : detection d'une encapsulation probable...

02/11/09 192.168.0.1->192.168.1.1 : ce couple de machines a echange 5 flux classes comme non WWW

02/11/09 192.168.0.1 : cette machine a emis 5 flux classes comme non WWW


02/11/09 192.168.0.1->192.168.1.1 : detection d'une encapsulation probable de session interactive (SSH ou TELNET).

02/11/09 192.168.0.1->192.168.1.1 : detection d'une encapsulation probable

#### Erreurs console

```
tcpdump: listening on eth1, link-type= EN10MB
(Ethernet), capture size 65535 bytes
1887 packets captured
1887 packets received by filter
0 packets dropped by kernel
```

**THALES**



**Base d'apprentissage**

- [Création](#)
- [Paramètres](#)

**Paramétrage de la classification**

- [Niveau paquet](#)
- [Niveau connexion](#)

**Paramétrage de l'IA**

- [Paramètres](#)

**Capture et analyse**

- [Capture](#)
- [Paramètres](#)

## Lancer une nouvelle capture...

Cette page permet de commencer une nouvelle écoute du trafic, et de visualiser en temps réel la classification des flux.

Timestamp	IP client	IP serveur	Port client	Port serveur	Classe	Conf.
02/11/09 17:11:39	192.168.0.1	192.168.1.1	60467	443	HTTPs	1
02/11/09 17:11:16	192.168.0.1	192.168.1.1	45886	51025	MAILs	0.95
02/11/09 17:11:59	192.168.0.1	192.168.1.1	49258	51110	MAILs	1
02/11/09 17:11:11	192.168.0.1	192.168.1.1	54515	51022	SSHs	0.75
02/11/09 17:11:06	192.168.0.1	192.168.1.1	54513	51022	SSHs	0.6
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60458	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60457	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60456	443	HTTPs	1
02/11/09 17:11:00	192.168.0.1	192.168.1.1	60455	443	HTTPs	1
02/11/09 17:11:22	192.168.0.1	192.168.1.1	35317	51110	MAILs	0.65

**Sortie console**


```
Mode de classification temps reel...
Lancement du demultiplexeur...
Lancement du script d'extraction des parametres...
Lancement du classificateur...
Utilisation du jeu de parametres : 2
Traitement d'un nouveau flux (1)...
Traitement d'un nouveau flux (2)...
Traitement d'un nouveau flux (3)...
Traitement d'un nouveau flux (4)...
Traitement d'un nouveau flux (5)...
Traitement d'un nouveau flux (6)...
Un flux classe niveau connexion : 6.0
```

**Rapport d'analyse**

02/11/09 17:11:19	192.168.0.1->192.168.1.1 : détection d'une encapsulation probable...
02/11/09 17:11:02	192.168.0.1->192.168.1.1 : détection d'une encapsulation probable...
02/11/09 17:11:20	192.168.0.1->192.168.1.1 : ce couple de machines a échangé 5 flux classés comme non WWW
02/11/09 17:11:20	192.168.0.1 : cette machine a émis 5 flux classés comme non WWW
02/11/09 17:11:20	192.168.0.1->192.168.1.1 : détection d'une encapsulation probable de session interactive (SSH ou TELNET)...
02/11/09	192.168.0.1->192.168.1.1 : détection d'une encapsulation probable

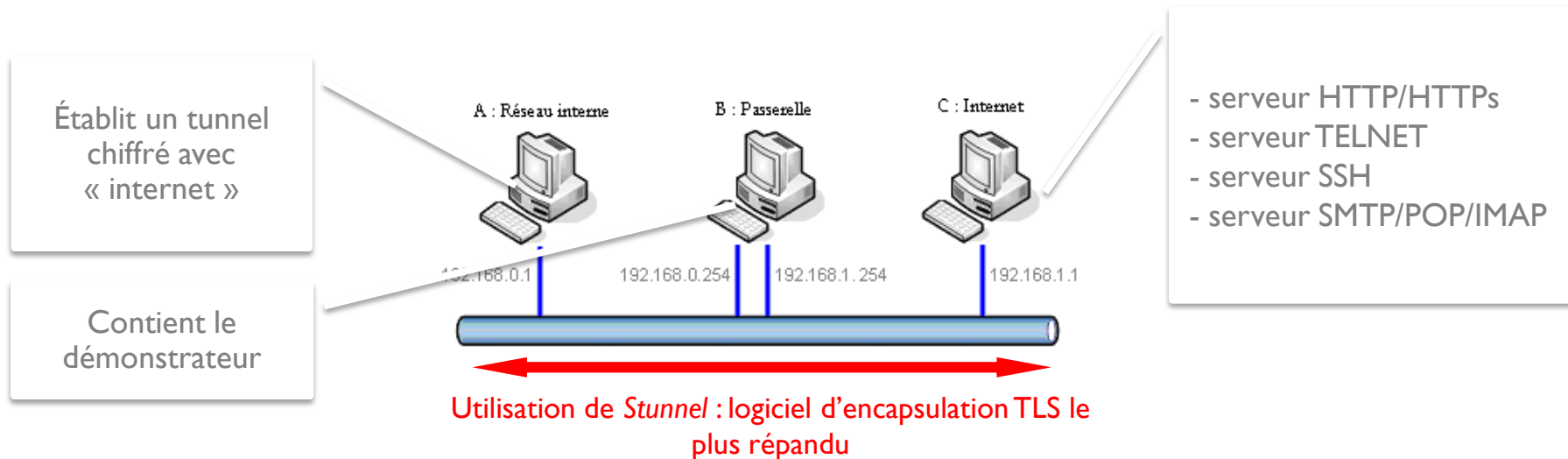
**Erreurs console**

```
tcpdump: listening on eth1, link-type EN10MB
(Ethernet), capture size 65535 bytes
1867 packets captured
1867 packets received by filter
0 packets dropped by kernel
```





## ■ Manipulations préliminaires sur plateforme réduite



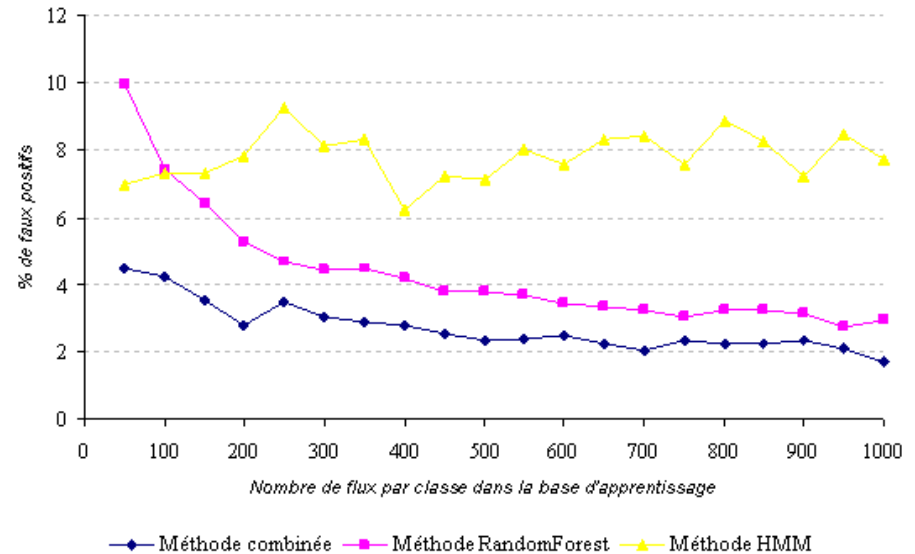
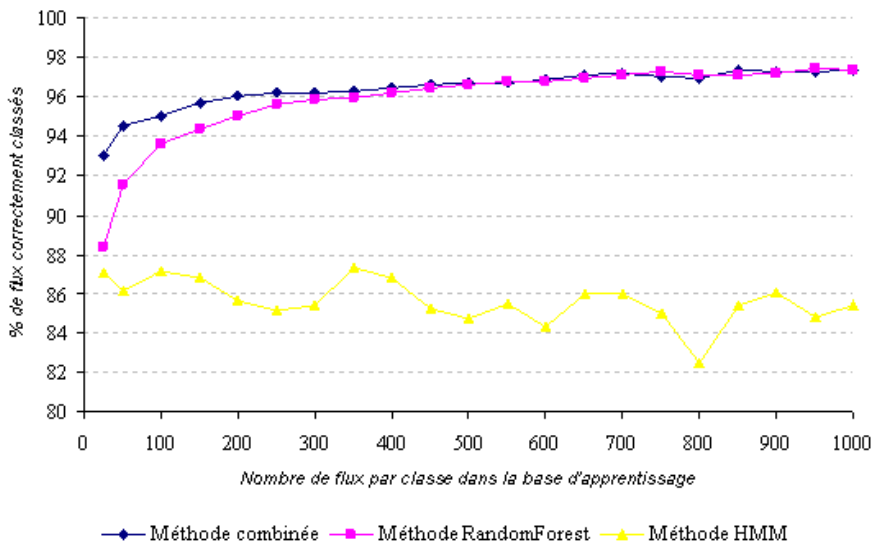
- Base d'apprentissage d'une vingtaine de flux pour chacun des 7 protocoles suivants : SSH, TELNET, HTTP, POP3, SMTP, SCP, IMAP
- Résultats : Taux de bonne identification du protocole par la passerelle proche de 100%

➤ Encourageant, mais pas suffisant : cette plateforme n'est pas représentative de la diversité des flux réels



- Base de données du groupe de travail MAWI
  - Composée de **flux réels**, capturés débuts 2009 sur une ligne transpacifique Japon/USA
  - Paquets **anonymisés** et données applicatives retirées
  - Reconstitution des flux, puis tri (élimination des protocoles sous-représentés, sous-échantillonnage de ceux sur-représentés)
  - Au total, près de **20 000 flux retenus**, appartenant à **9 protocoles différents**

<b>Protocole</b>	<b>HTTP</b>	<b>HTTPs</b>	<b>SSH</b>	<b>SMTP</b>	<b>DNS</b>	<b>FTP</b>	<b>Active Directory</b>	<b>POP3s</b>	<b>Active NetSteward</b>
<b>Nombre de flux</b>	2500	2500	2500	2500	2500	2500	1069	1503	1611



- Résultats concluants : **taux de bonne classification > 96%**
- Importance d'obtenir une base d'apprentissage de taille raisonnable
- Avantages de la méthode combinée :
  - **réduction de 35% du nombre de faux positifs** (par rapport à *RandomForest*)
  - meilleurs résultats sur les petites bases d'apprentissage



## ■ Performances temporelles

	<i>RandomForest</i> seul	Banc de HMM seul	CASTAFIOR
<i>Durée de la phase d'apprentissage (2500 flux)</i>	≈ 11 s.	≈ 41 s.	≈ 45s
<i>Durée moyenne de classification d'un flux</i>	≈ 0,22 ms.	≈ 0,21 μs.	≈ 0.5 ms.

- NB : performances calculées sur une Debian Lenny, dualcore Pentium 4 à 3.06Gz, JVM Open-jdk 1.6.0
- Le code de classification est en Java, et non optimisé





- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan



- De **bons résultats** qualitatifs et quantitatifs...
  - Une majorité de tunnels illégitimes peuvent être détectés
  - La combinaison des deux méthodes d'analyse améliore les performances
  - Le temps de calcul est raisonnable compte tenu de l'implémentation
  
- Mais...
  - Un taux de **fausses alarmes** encore trop élevé
  - Une constitution de la **base d'apprentissage** problématique (elle doit être suffisamment représentative)
  - Un comportement aléatoire d'un tel outil face à un flux d'un protocole inconnu
  - Un certain nombre de travaux restent à faire
    - Etudier les possibilités de contournement
    - Analyser les erreurs fréquentes, etc.

**Un complément prometteur aux outils actuels, à intégrer comme source d'information supplémentaire dans une architecture de sécurité plus vaste**





# BACKUPS



- Identification des flux par des propriétés statistiques : **sujet riche dans la littérature** académique depuis 2005
  - Quelques brevets déposés (Alcatel-Lucent début 2009)
- Les études sont rarement orientées sécurité
  - Plutôt QoS, modélisation réseau, etc.
- Paramètres généralement retenus
  - Moyenne, variance, max et min pour les temps d'inter arrivées de paquets et les tailles des paquets
- La plupart des **méthodes d'apprentissage** (supervisées ou non) ont été appliquées à ce sujet
- Des **résultats prometteurs...**
  - Taux d'identification du protocole généralement supérieur à 90%
- **...mais impossibles à comparer** d'une étude à l'autre
  - Base de données et méthodologie différentes pour chaque étude

# ANNEXE B : Matrice de confusion

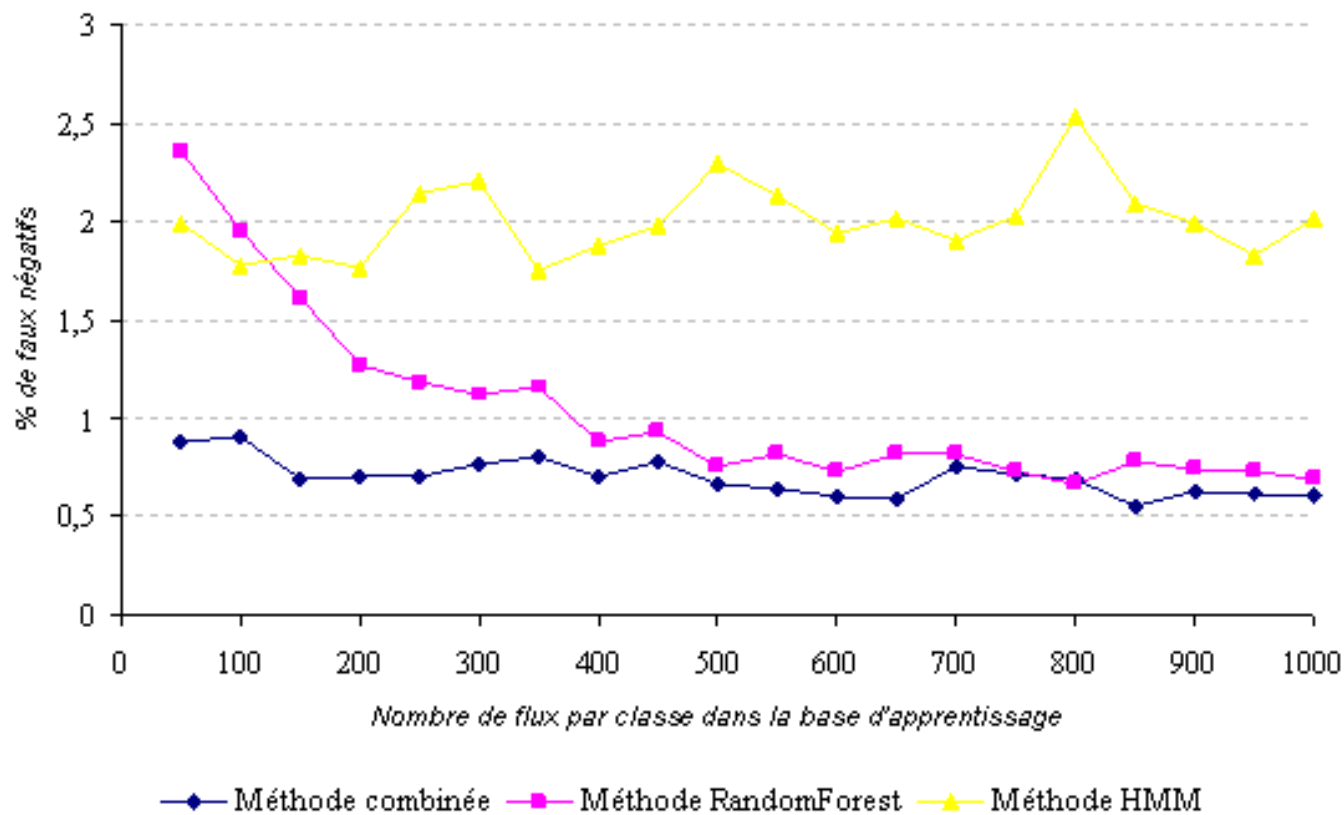


<i>HTTP</i>	<i>HTTPs</i>	<i>SSH</i>	<i>SMTP</i>	<i>DNS</i>	<i>FTP</i>	<i>Act.Dir.</i>	<i>POP3s</i>	<i>NetStew</i>	
<b>94.88</b>	2.84	0.0	0.6	0.2	0.16	0.12	0.52	0.68	<i>HTTP</i>
1.72	<b>94.44</b>	0.0	1.8	0.12	0.4	0.44	0.92	0.16	<i>HTTPs</i>
0.0	0.0	<b>99.56</b>	0.2	0.0	0.0	0.08	0.16	0.0	<i>SSH</i>
0.0	1.72	0.0	<b>89.64</b>	0.08	3.32	4.0	0.4	0.66	<i>SMTP</i>
0.0	0.0	0.0	0.24	<b>99.68</b>	0.04	0.04	0.0	0.0	<i>DNS</i>
0.12	0.64	0.0	2.64	0.24	<b>94.96</b>	0.76	0.4	0.24	<i>FTP</i>
0.0	0.28	0.0	0.56	0.0	0.0	<b>99.16</b>	0.0	0.0	<i>ActiveDirectory</i>
0.67	1.06	0.0	0.67	0.0	0.06	0.0	<b>97.54</b>	0.0	<i>POP3s</i>
0.74	0.0	0.0	0.0	0.07	0.0	0.0	0.0	<b>99.19</b>	<i>NetSteward</i>

Matrice de confusion obtenue par application de CASTAFIOR à la base de données MAWI



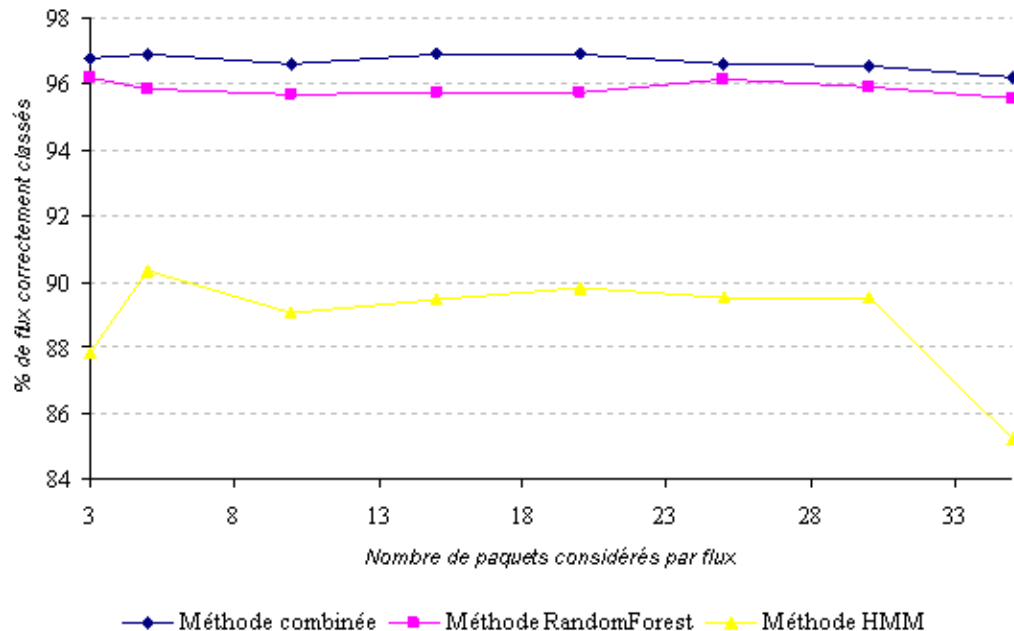
- **Comparaison de 5 méthodes** de classification par apprentissage ;
- **Combinaison de deux méthode de classification** à des niveaux complémentaires : « macroscopique » et « microscopique » ;
- **Conception et développement d'un outil complet** de détection des tunnels, implémentant ces deux méthodes ainsi qu'une « **intelligence artificielle** » élémentaire pour limiter les faux positifs ;
- Présentation des **résultats** de l'application de notre outil à une base de données de flux réels de grande taille.







## ■ Nombre de paquets pris en compte avant la classification



- Étonnamment, performances peu affectées !
- Comportements distinguables dès le début des connexions (pour nos protocoles)
- MAIS : si petit nombre de paquets pris en compte, contournement trivial possible !