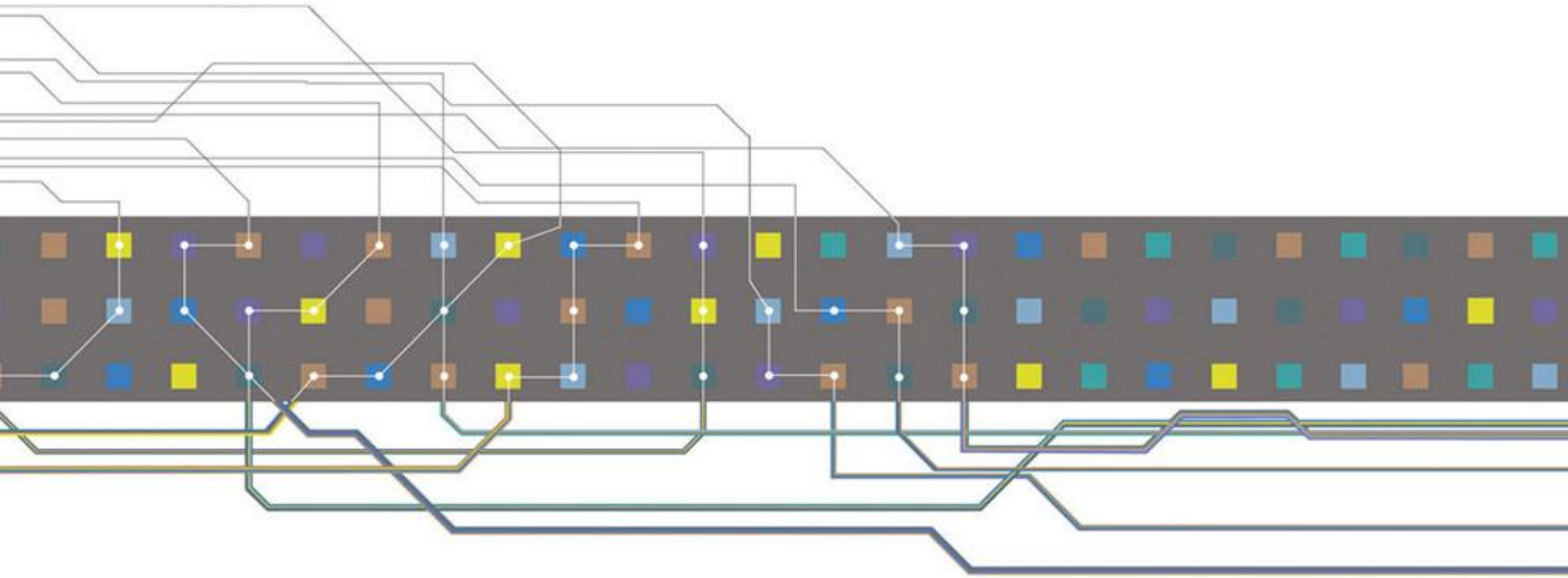
THALES



CASTAFIOR

Détection Automatique de Tunnels Illégitimes par Analyse Statistique

Fabien ALLARD – Mathieu MOREL

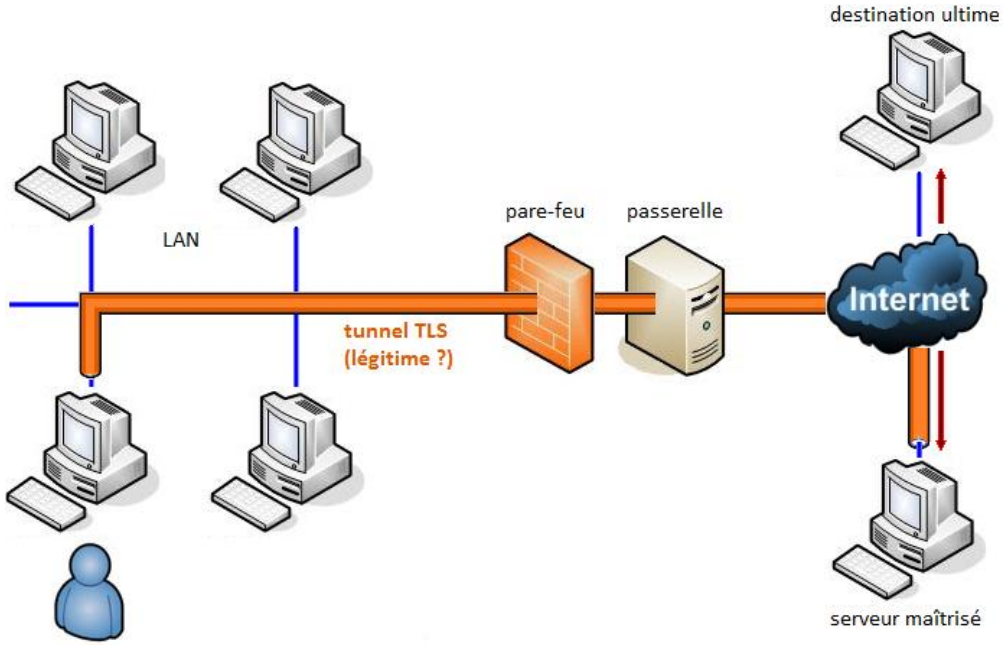
Plan de la présentation

- 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

Plan de la présentation

- 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

Cadre et problématique



- 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.)

Problématique

- Les méthodes de protection possibles
  - Cryptanaliser 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.

Les enjeux

- 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.

Plan de la présentation

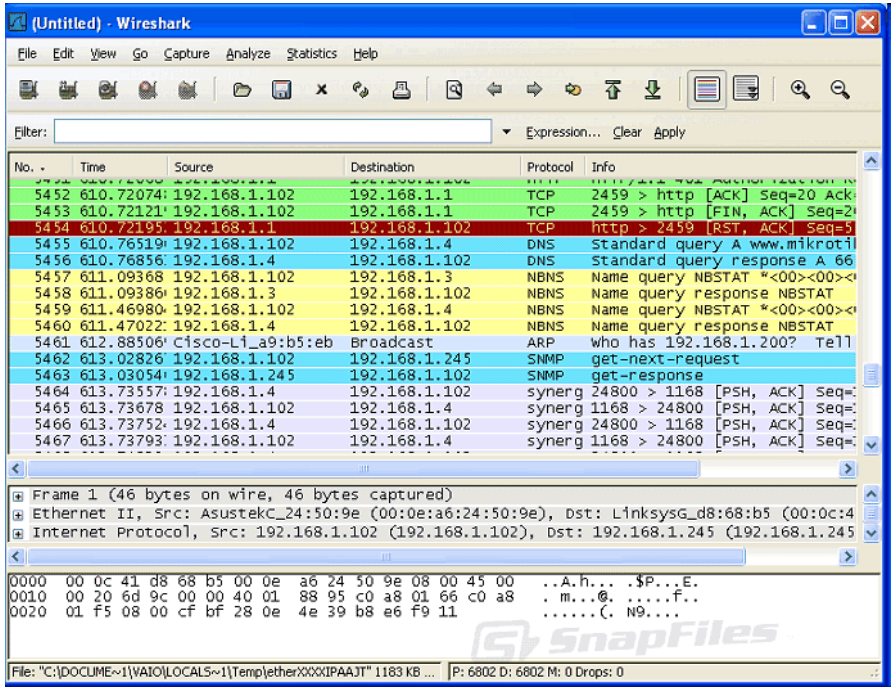
- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan

Fonctionnement d'une classification par apprentissage (1)

- 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

Fonctionnement d'une classification par apprentissage (2)

Constitution d'une base de données d'apprentissage

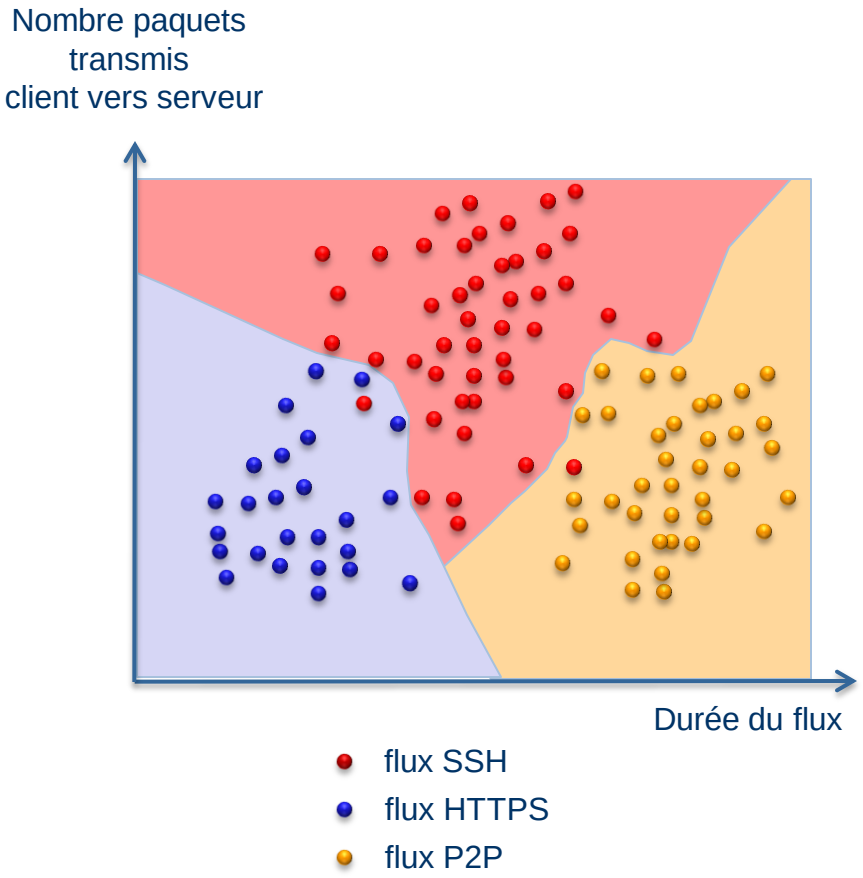


Fonctionnement d'une classification par apprentissage (2)

Constitution d'une base de données d'apprentissage

Projection de chaque observation sur l'espace des paramètres

Construction d'un modèle



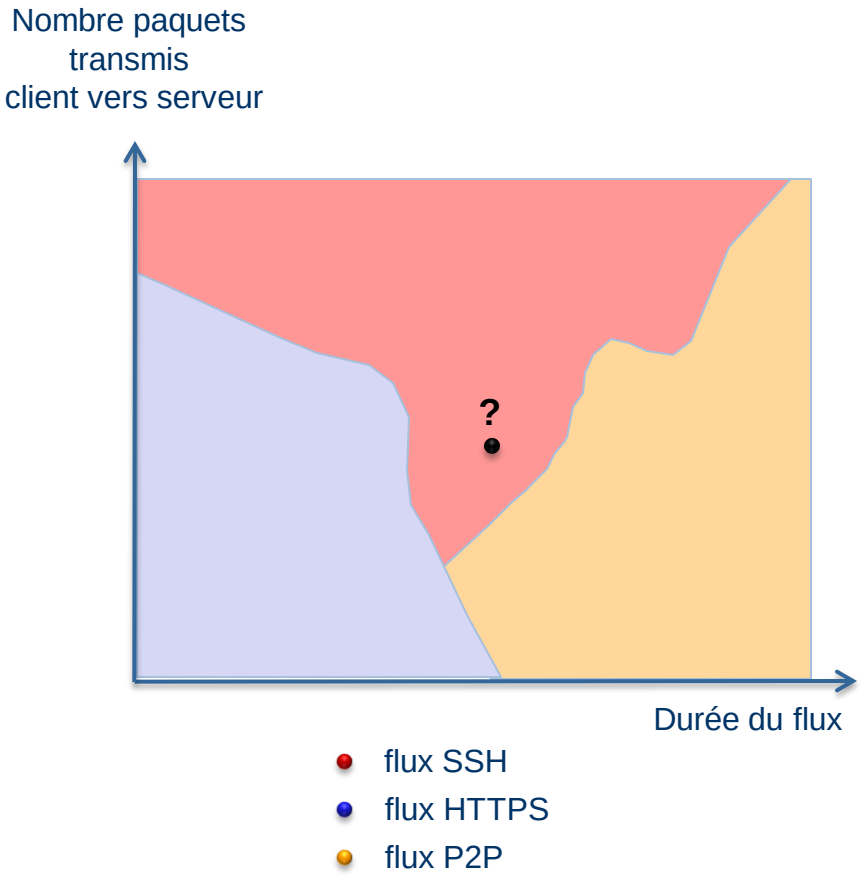
Fonctionnement d'une classification par apprentissage (2)

Constitution d'une base de données d'apprentissage

Projection de chaque observation sur l'espace des paramètres

Construction d'un modèle

Classification des flux à la volée à partir du modèle



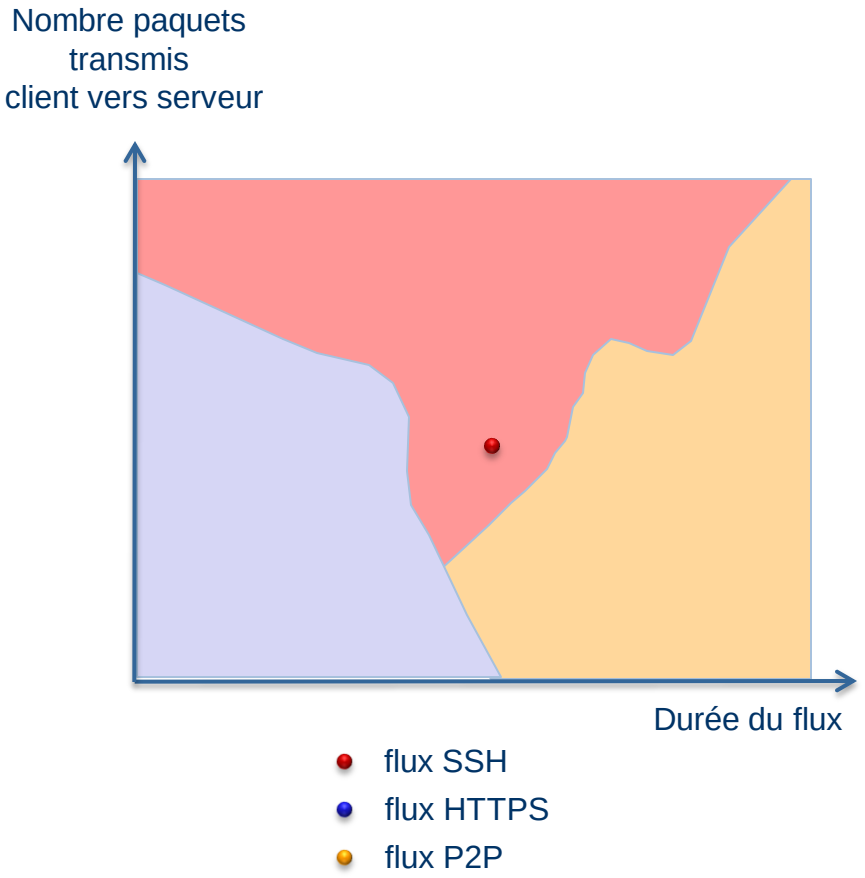
Fonctionnement d'une classification par apprentissage (2)

Constitution d'une base de données d'apprentissage

Projection de chaque observation sur l'espace des paramètres

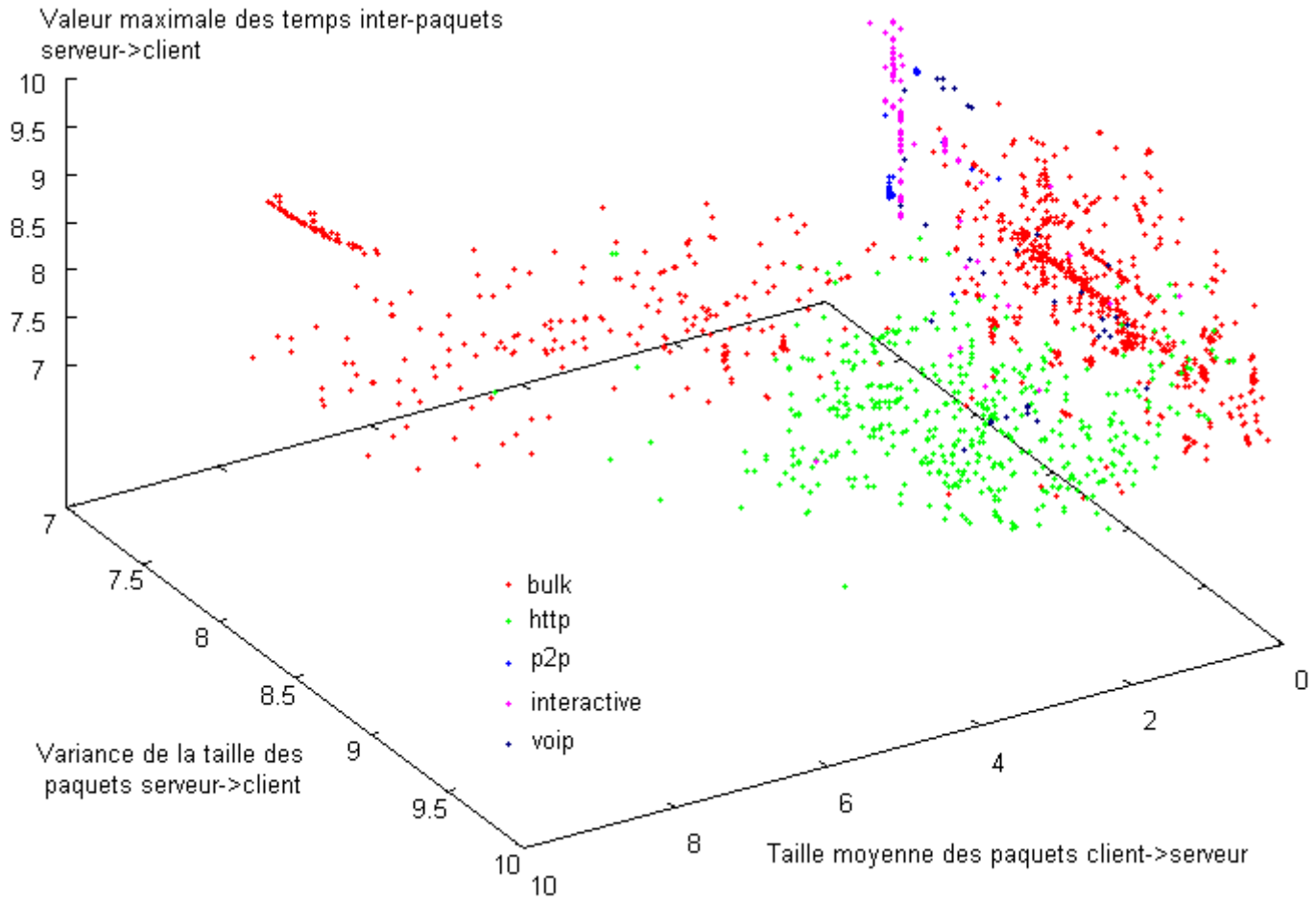
Construction d'un modèle

Classification des flux à la volée à partir du modèle



THALES

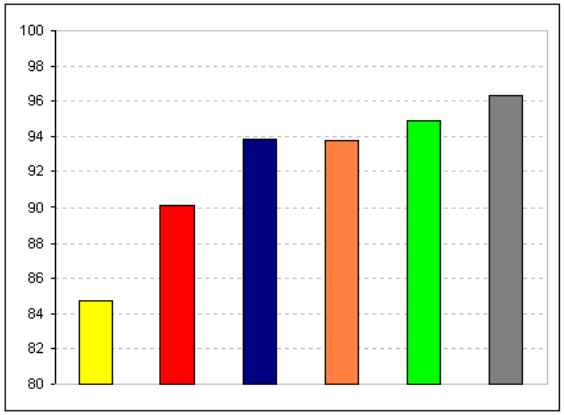
Fonctionnement d'une classification par apprentissage (3)



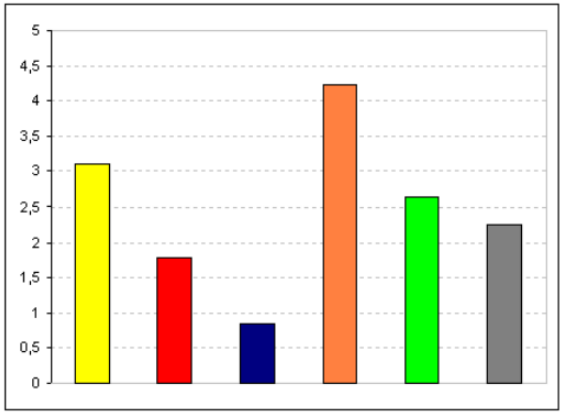
Projection de quelques flux réels sur trois paramètres utilisés pour la classification

Choix d'une méthode de classification « macroscopique »

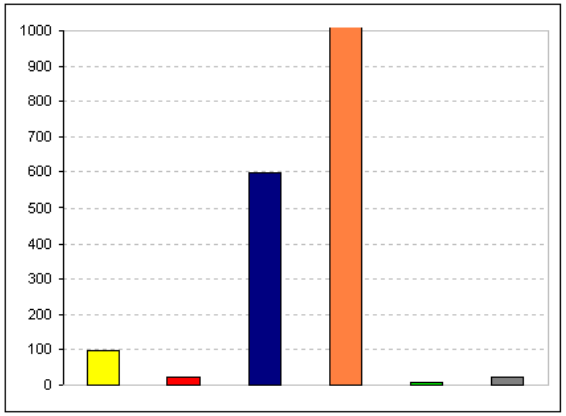
- Utilisation d'une base de données publique de 20 000 flux clairs
- Détermination des 10 paramètres à extraire des flux
  - Algorithmes de sélection automatique des paramètres les plus discriminants et les moins corrélés
- A partir de l'étude bibliographique, applications de 6 méthodes d'apprentissage
  - 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



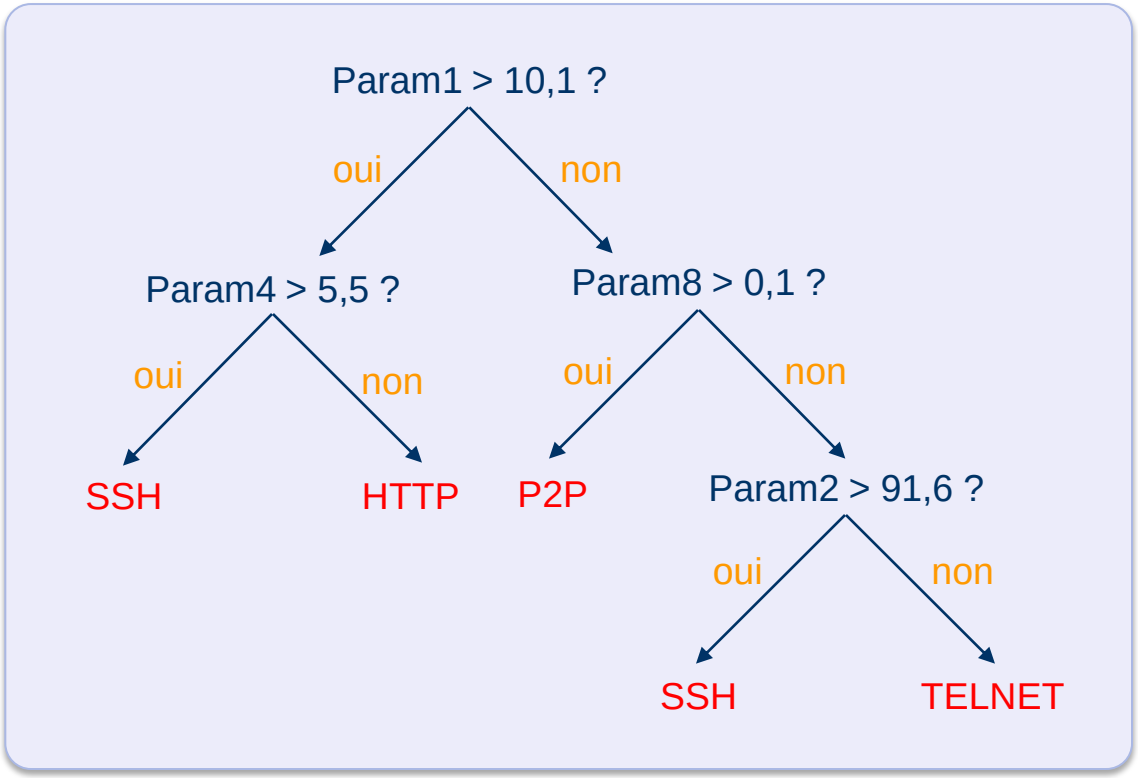
Taux de faux positifs
(HTTP mal classés)



Temps de classification (en µs)

La méthode *RandomForest*

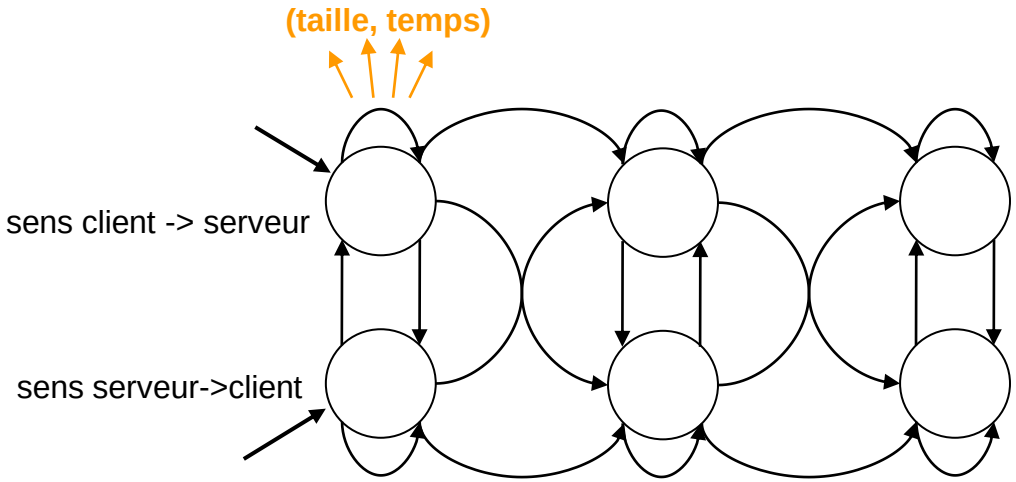
- Qu'est-ce qu'un arbre de décision ?



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

Une approche « microscopique » : Markov caché

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

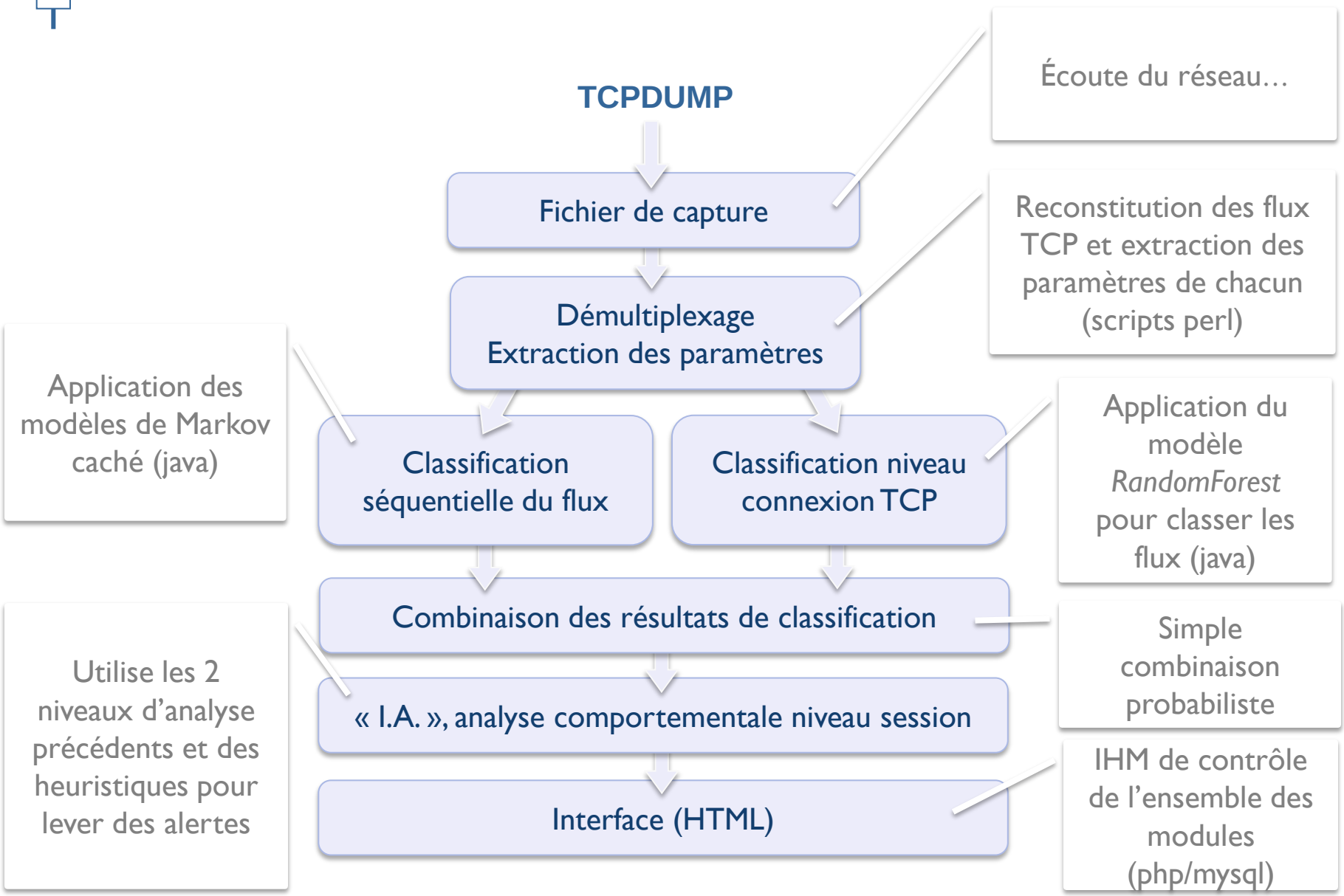


- Apprentissage (à partir de la base d'apprentissage)
  - Calcul des probabilités de transition et d'émission par l'algorithme de Baum Welsh
- Classification
  - On calcule sa probabilité d'émission par chacun des HMM (algorithme de Viterbi)
  - On retient le HMM ayant donné la probabilité maximale

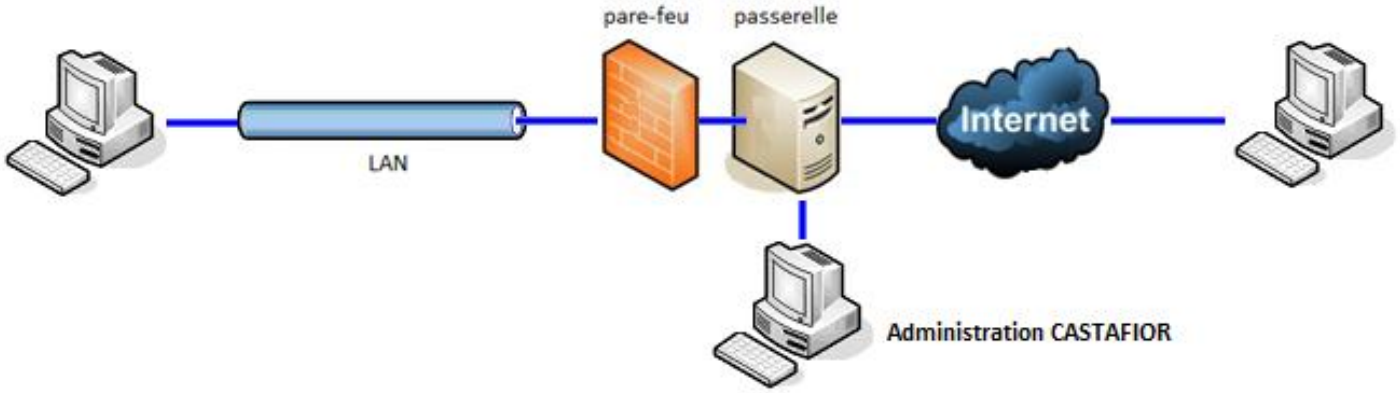
Plan de la présentation

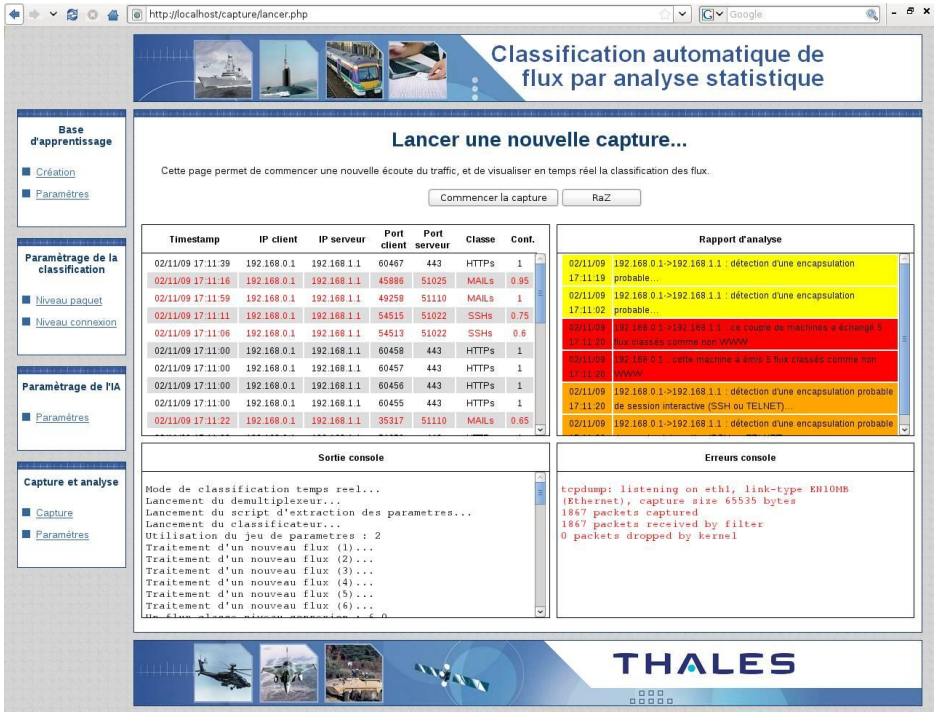
- Contexte, problématique
- Étude théorique
  - Principe de fonctionnement d'une méthode de classification
  - Analyse « macroscopique » : *RandomForest*
  - Analyse « microscopique » : *Modèles de Markov cachés*
- Expérimentations
  - Fonctionnement de CASTAFIOR
  - Résultats obtenus
- Bilan

Architecture de CASTAFIOR

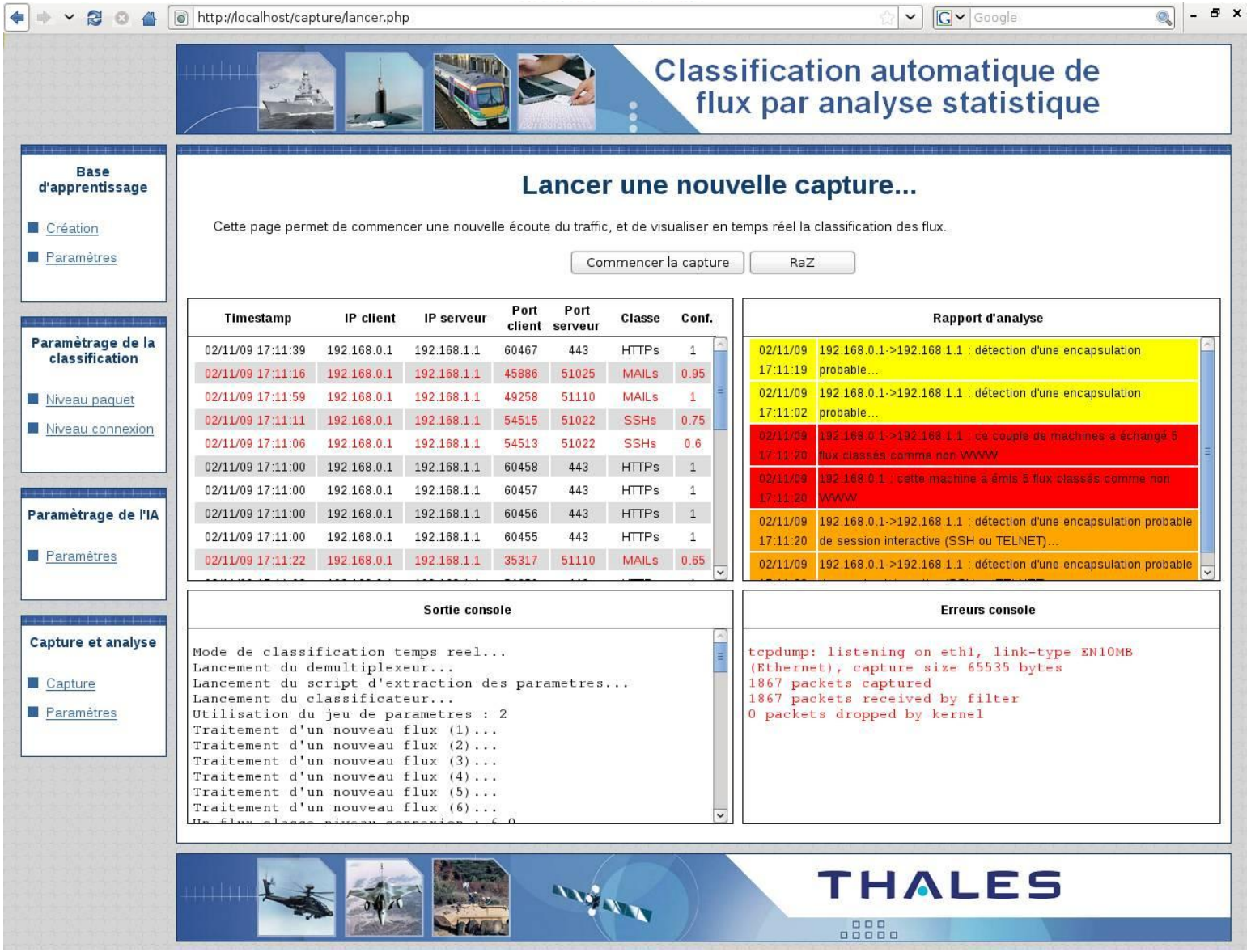


Utilisation de CASTAFIOR





Utilisation de CASTAFIOR

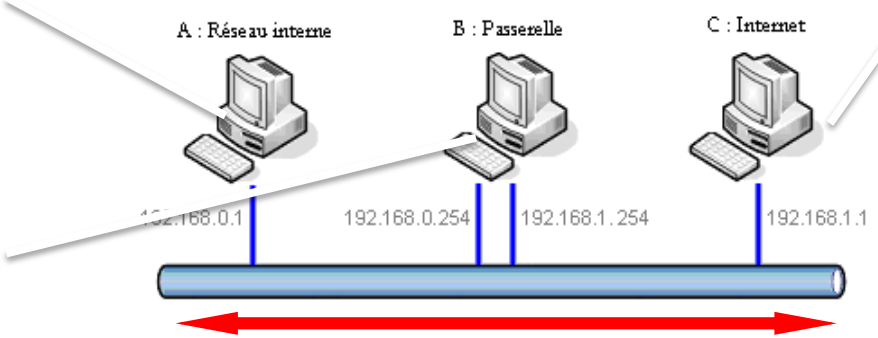


Expérimentations 1 : plateforme réduite

- Manipulations préliminaires sur plateforme réduite

Établit un tunnel chiffré avec « internet »

Contient le démonstrateur



- serveur HTTP/HTTPs
- serveur TELNET
- serveur SSH
- serveur SMTP/POP/IMAP

Utilisation de *Stunnel* : logiciel d'encapsulation TLS le plus répandu

- 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%

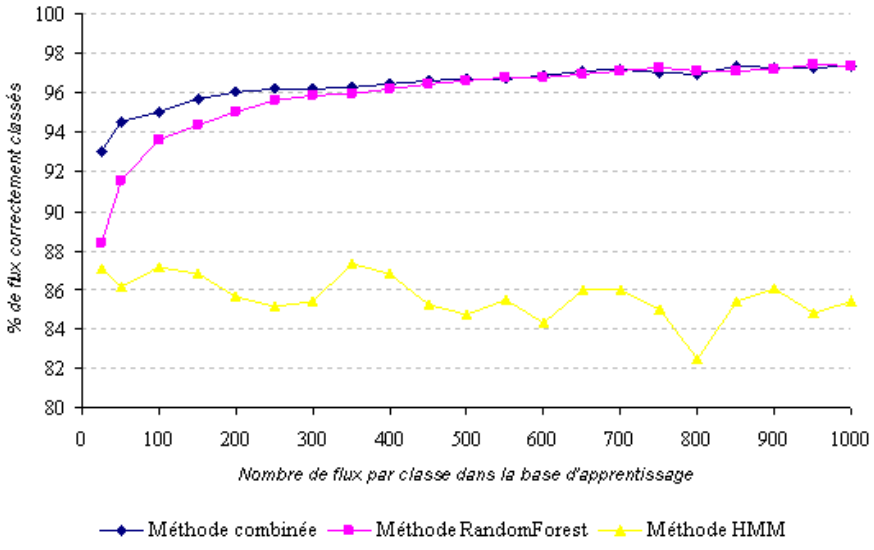
➢ Encourageant, mais pas suffisant : cette plateforme n'est pas représentative de la diversité des flux réels

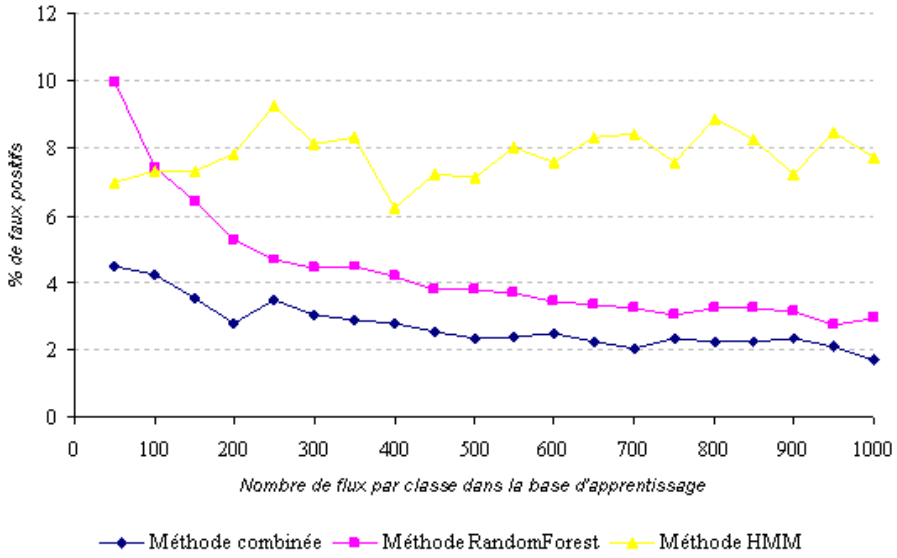
Expérimentations 2 : flux réels

- 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-échantillonage de ceux sur-représentés)
  - Au total, près de 20 000 flux retenus, appartenant à 9 protocoles différents

Protocole	**HTTP**	**HTTPs**	**SSH**	**SMTP**	**DNS**	**FTP**	**Active Directory**	**POP3s**	**Active NetSteward**
Nombre de flux	2500	2500	2500	2500	2500	2500	1069	1503	1611

Résultats obtenus





- 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

Résultats obtenus

- Performances temporelles

	***RandomForest* seul**	**Banc de HMM seul**	**CASTAFIOR**
Durée de la phase d'apprentissage (2500 flux)	≈ 11 s.	≈ 41 s.	**≈ 45s**
Durée moyenne de classification d'un flux	≈ 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

Plan de la présentation

- 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

Questions

BACKUPS

ANNEXE A : L'état de l'art

- 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

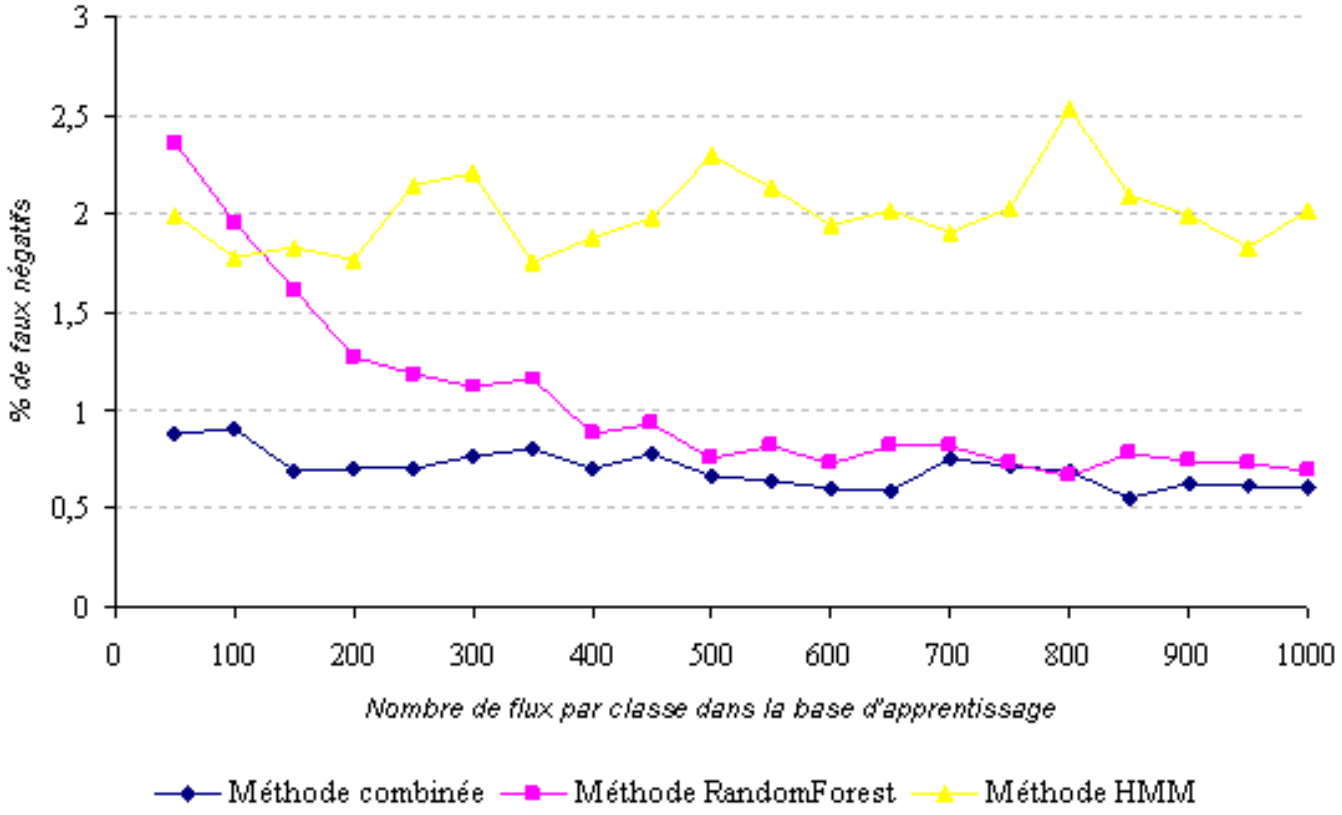
HTTP	*HTTPs*	*SSH*	*SMTP*	*DNS*	*FTP*	*Act.Dir.*	*POP3s*	*NetStew*	
94.88	2.84	0.0	0.6	0.2	0.16	0.12	0.52	0.68	*HTTP*
1.72	**94.44**	0.0	1.8	0.12	0.4	0.44	0.92	0.16	*HTTPs*
0.0	0.0	**99.56**	0.2	0.0	0.0	0.08	0.16	0.0	*SSH*
0.0	1.72	0.0	**89.64**	0.08	3.32	4.0	0.4	0.66	*SMTP*
0.0	0.0	0.0	0.24	**99.68**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

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

ANNEXE C : Notre contribution

- 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.

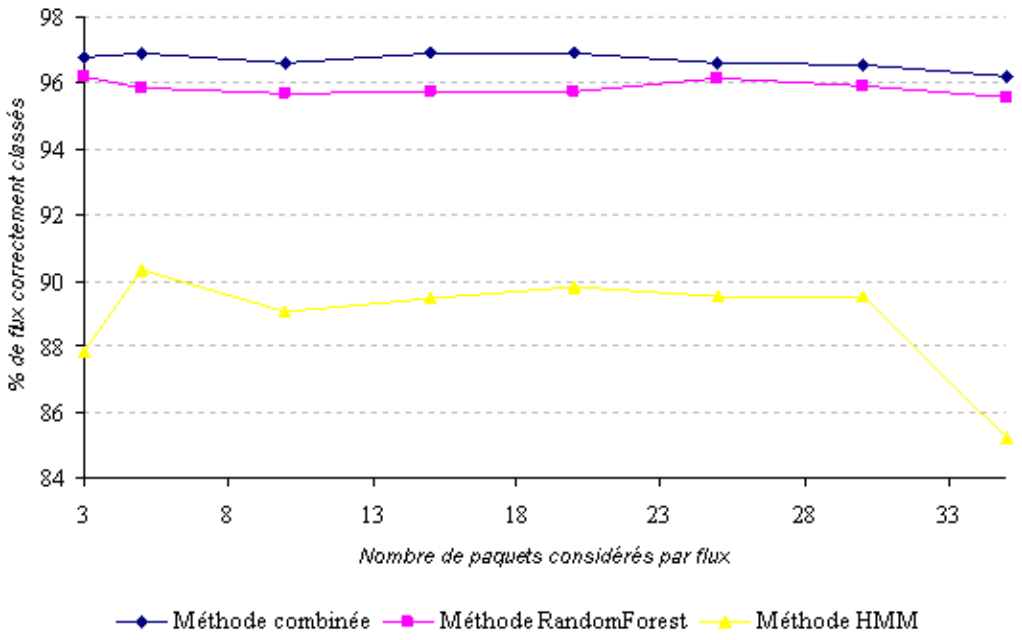
ANNEXE D : Faux négatifs



THALES

ANNEXE E : Influence du nombre de paquets

- Nombre de paquets pris en compte avant la classification



- Etonnamment, 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 !