



# Profils Propres Pour la Détection d'Intrusion

*{Yacine.Bouzida,Sylvain.Gombault}@enst-bretagne.fr*



# ■ Plan

---

- Introduction
- Méthodes de Détection d'Intrusion
- Présentation de notre Méthode (les Profils Propres)
  - L'Analyse en Composantes Principales
  - Les Différentes Etapes de la Méthode
  - Expérimentation (Navigation sur le Web)
- Conclusion et Perspectives

# ■ Introduction

---



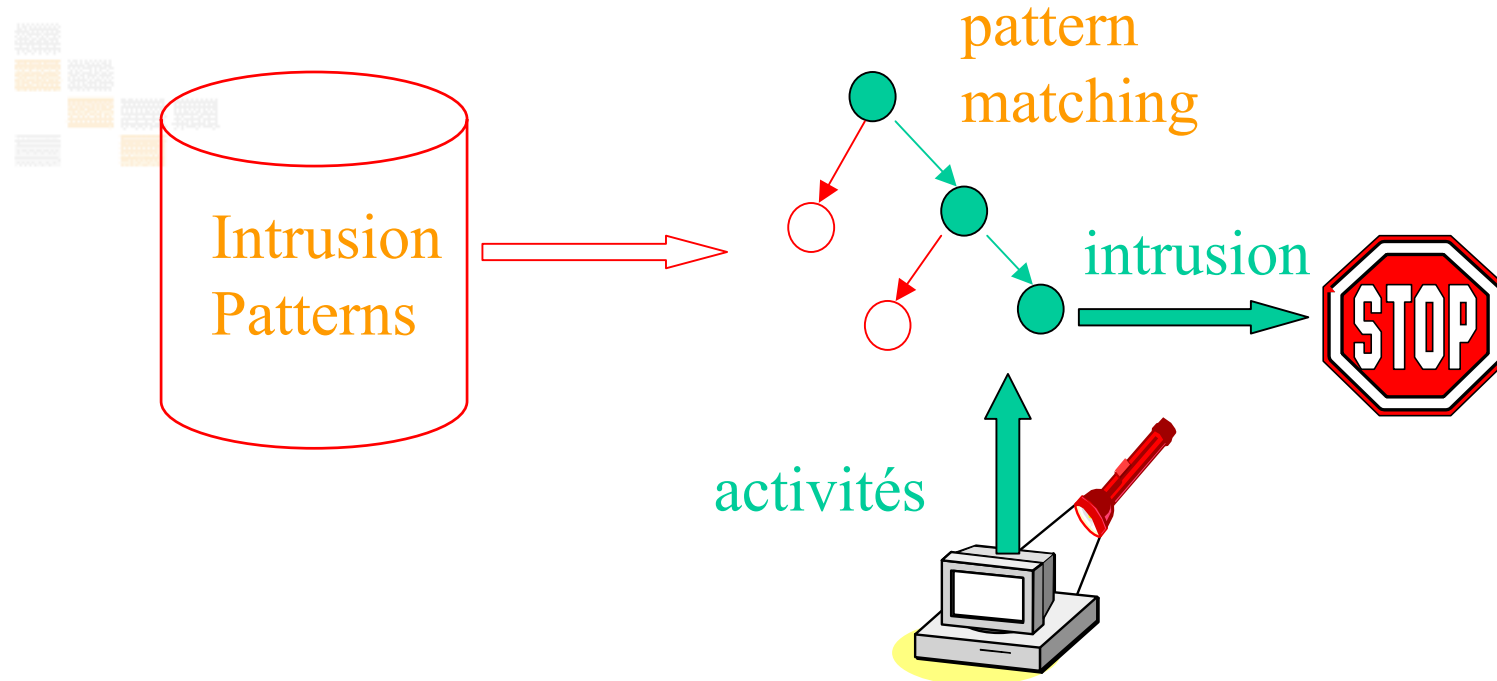
## ■ Intrusion

- (Objectif) Toute violation de la politique de sécurité d'un système d'information

## ■ Méthode

- 1 - Définir une politique de sécurité et la mettre en œuvre
- 2 - Surveiller afin d'assurer le respect de la politique

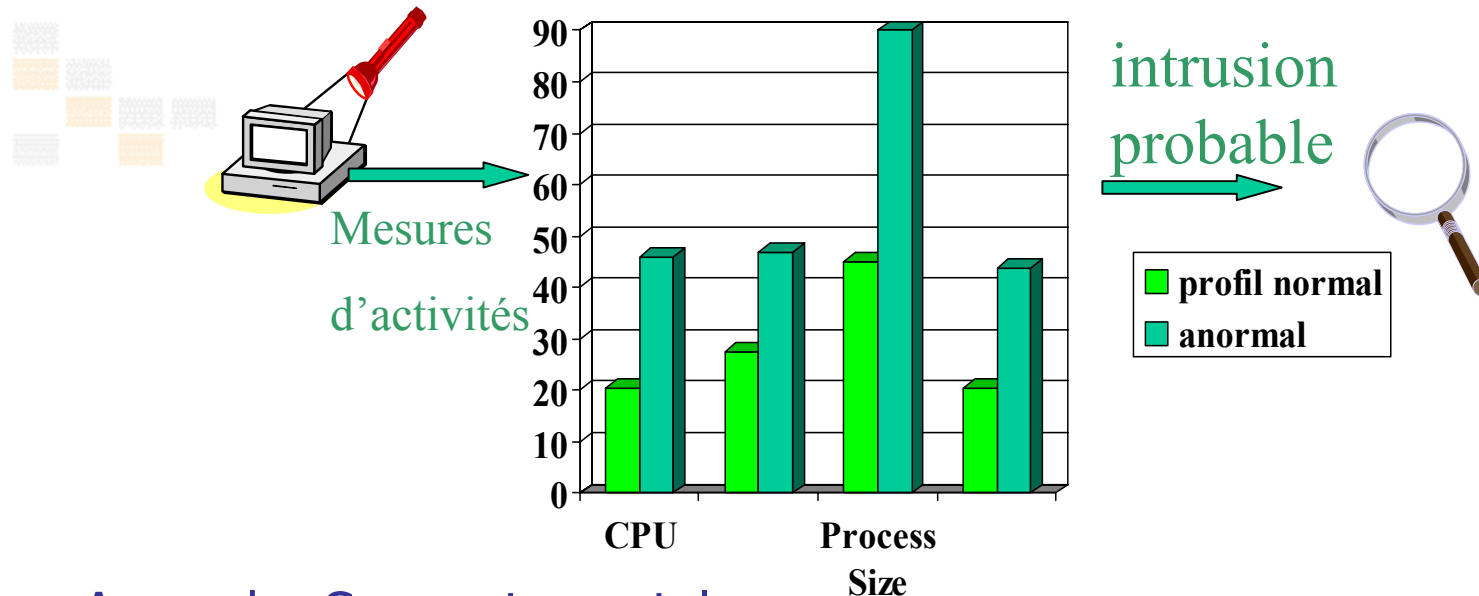
# ■ Approches de Détection



## ■ Approche Par Scénario

- Cherche à répondre à la question : "le comportement actuel de l'utilisateur correspond-il à un comportement intrusif connu? "
- Nécessite donc de construire une base de données d'attaques ou tout au moins "d'actions litigieuses"
- **Ne peut pas détecter des attaques non connues (nouvelles)**

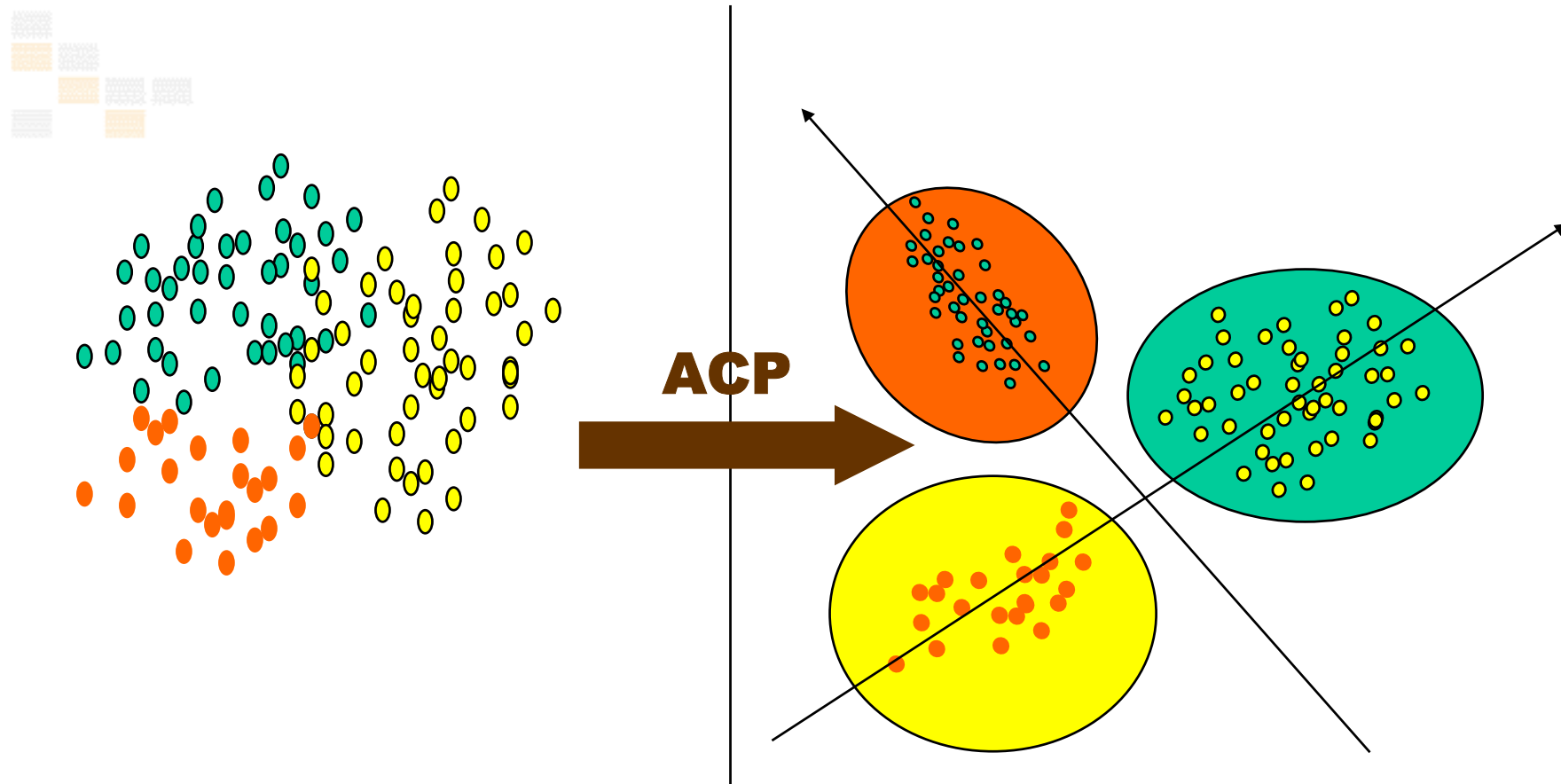
# ■ Approches de Détection



## ■ Approche Comportementale

- Proposée par Anderson dès 1980 et reprise par Denning en 1987.
- Taux de fausses alarmes très élevé - les anomalies peuvent être juste de nouvelles activités normales.
- Mais reste toujours le moyen de déceler d'éventuelles usurpations d'identité provenant d'attaque inconnue.

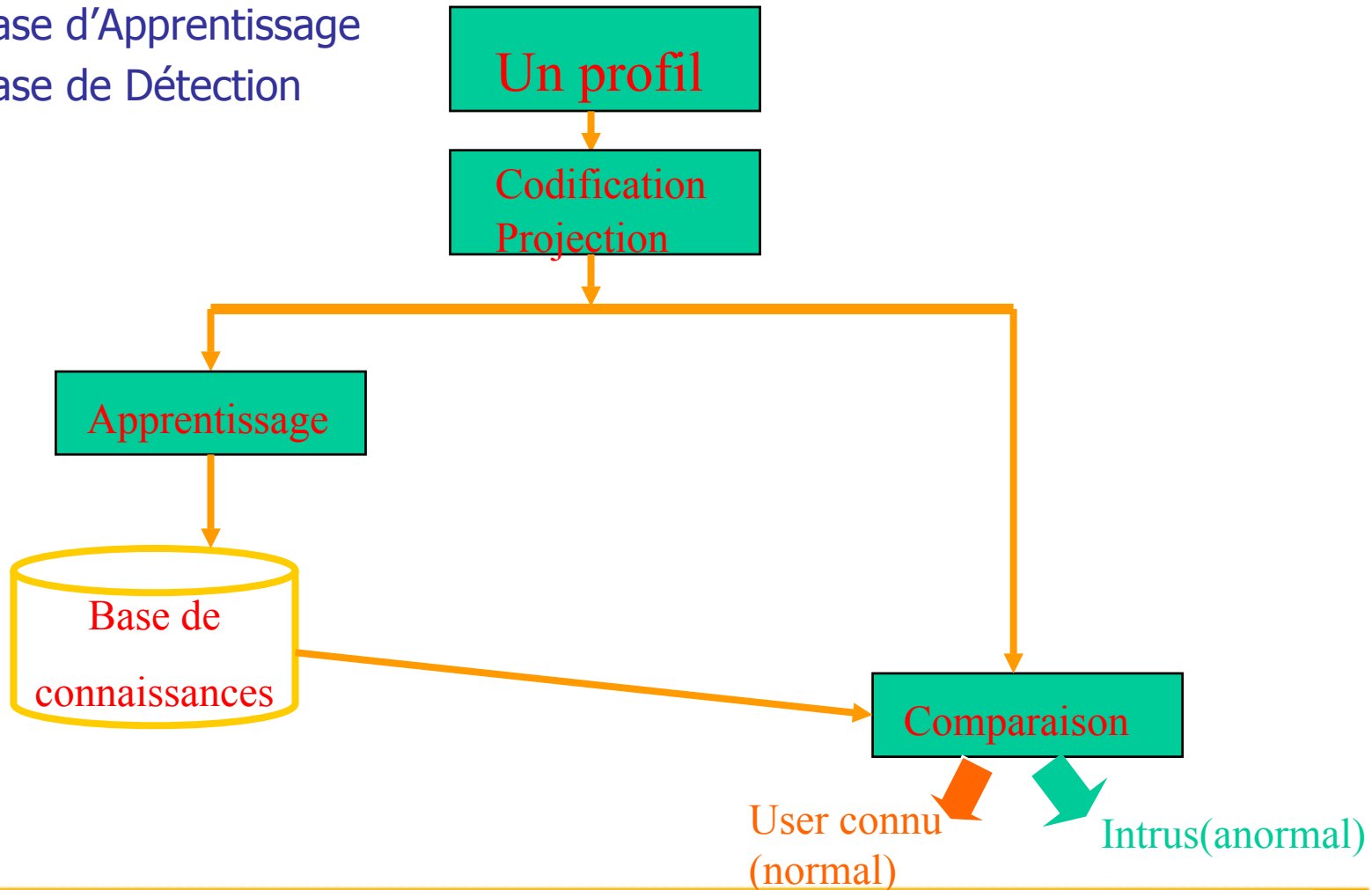
## ■ Les Profils Propres pour une approche comportementale



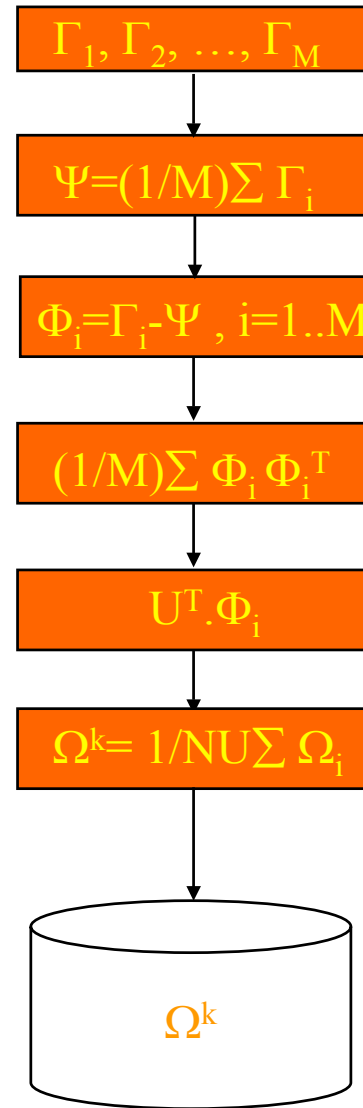
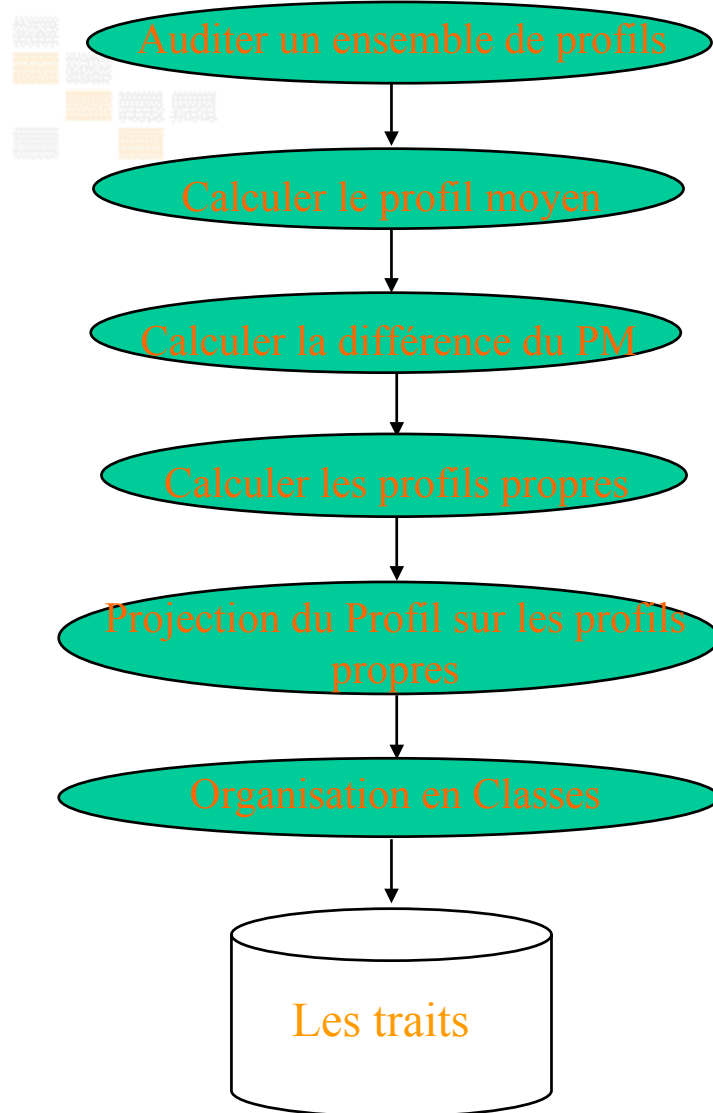
- Réduction d'un système complexe en un plus petit nombre de dimensions
- Permet la classification.

# ■ Présentation de la Méthode

- Le Système comporte deux phases
  - Phase d'Apprentissage
  - Phase de Détection



# Phase d'Apprentissage



# Phase de Détection

---



- Auditer un nouveau profil:  $\Gamma_i$
- Déterminer son Profil Caricature (Différence):  $\Phi_i = \Gamma_i - \Psi$
- Le projeter dans l'espace des profils propres :  $\Omega_i = \mathbf{U}^T \cdot \Phi_i$
- Définir la classe k qui minimise la distance  $\mathbf{d} = \|\Omega_i - \Omega^k\|_2$

**Si**  $d < \text{seuil } \theta_k$  **Alors**

- Le profil correspond au k<sup>ième</sup> utilisateur

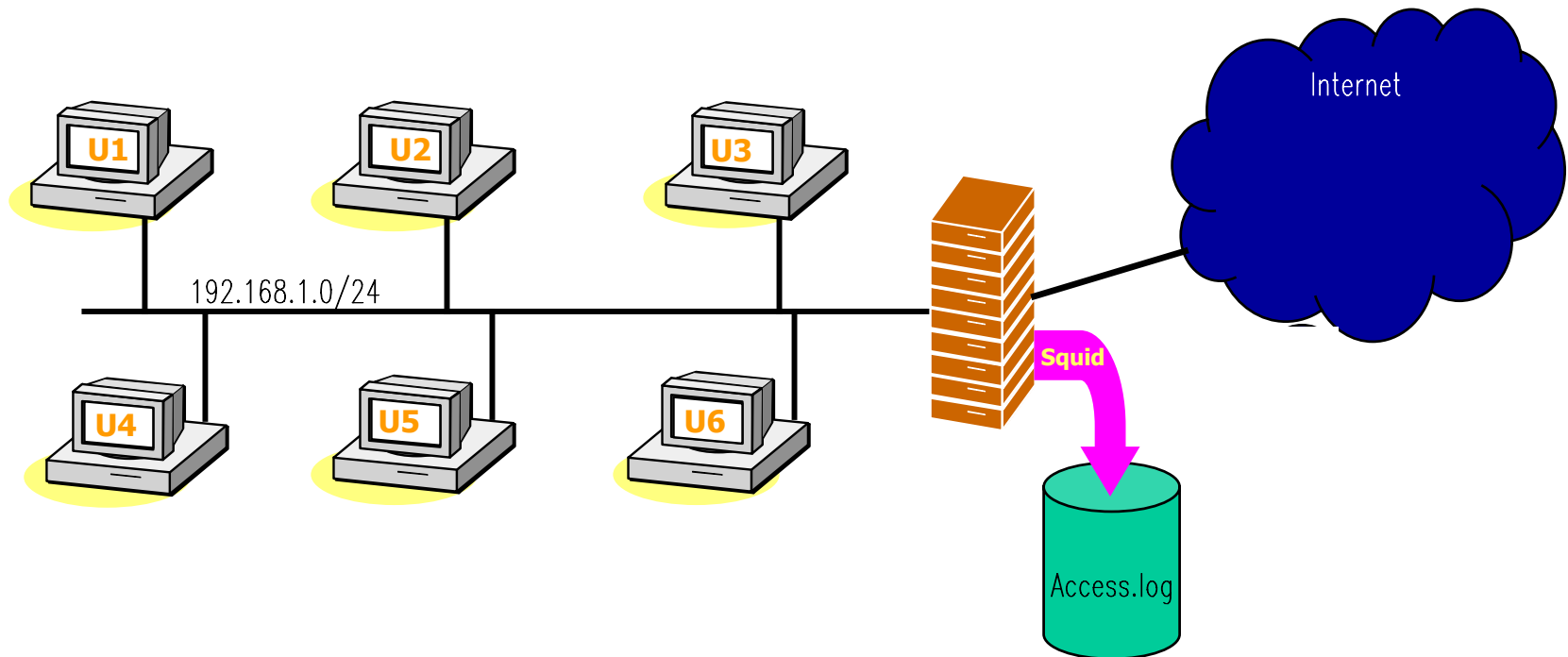
**Sinon**

- Comportement anormal « intrus »

**FSi**

# ■ Expérimentation

- Buts: Identifier l'utilisateur d'un réseau de stations de travail  
Détecter une usurpation de poste de travail,  
⇒ Utiliser une approche comportementale et l'audit du proxy web.
  - Squid



# Audit squid



**Squid**

*access.log*

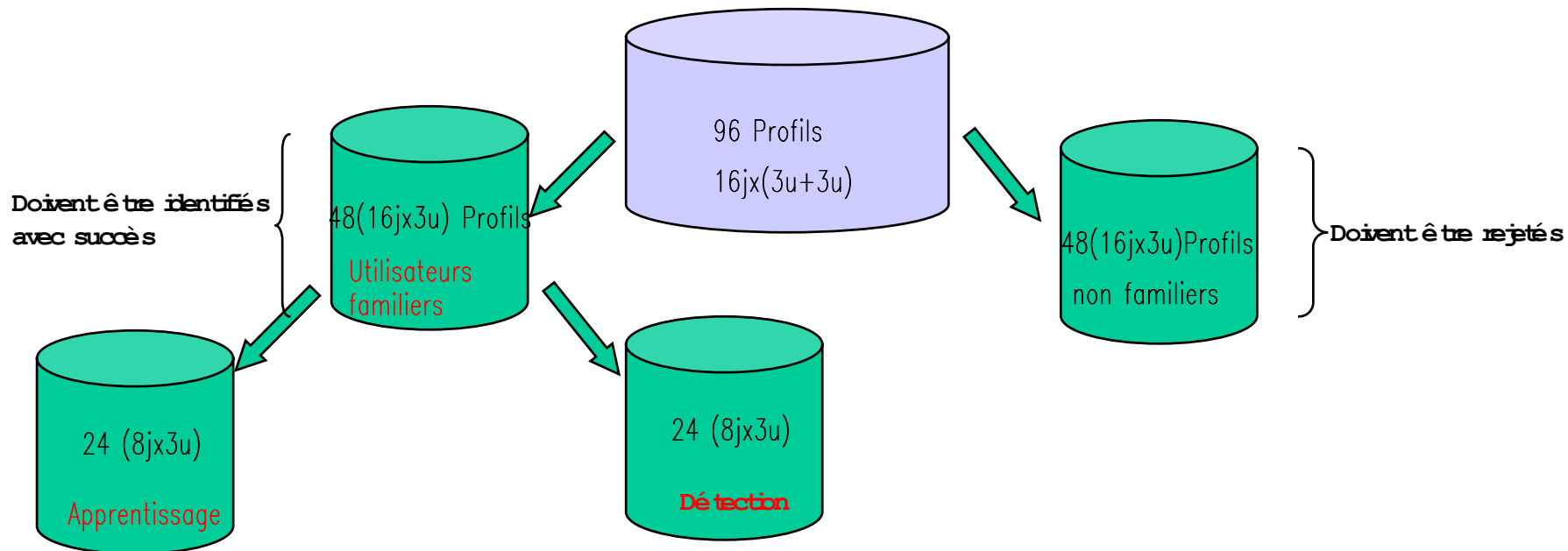
```
10:35:41.5 ... 192.168.1.2 ... http://www.google.fr/ ...  
10:35:42.2 ... 192.168.1.2 ... http://www.google.fr/ ...  
10:35:45.6 ... 192.168.1.2 ... http://www.google.fr/ ...  
10:35:45.6 ... 192.168.1.2 ... http://www.tcpdump.org/ ...  
...
```

**Profil**

<i>www.google.fr</i>	3
<i>www.tcpdump.org</i>	1

# ■ Application

- Audit de 6 utilisateurs. Chaque élément du vecteur profil représente le nombre de fois qu'une page web correspondant à cet élément a été visitée pendant la session d'audit,
- 1 profil d'un utilisateur  $\Rightarrow$  1 journée d'audit
- 1 mois d'audit (16 jours)
- 96 profils utilisés pour les tests (soit 16 profils par utilisateur).



# ■ Application

---

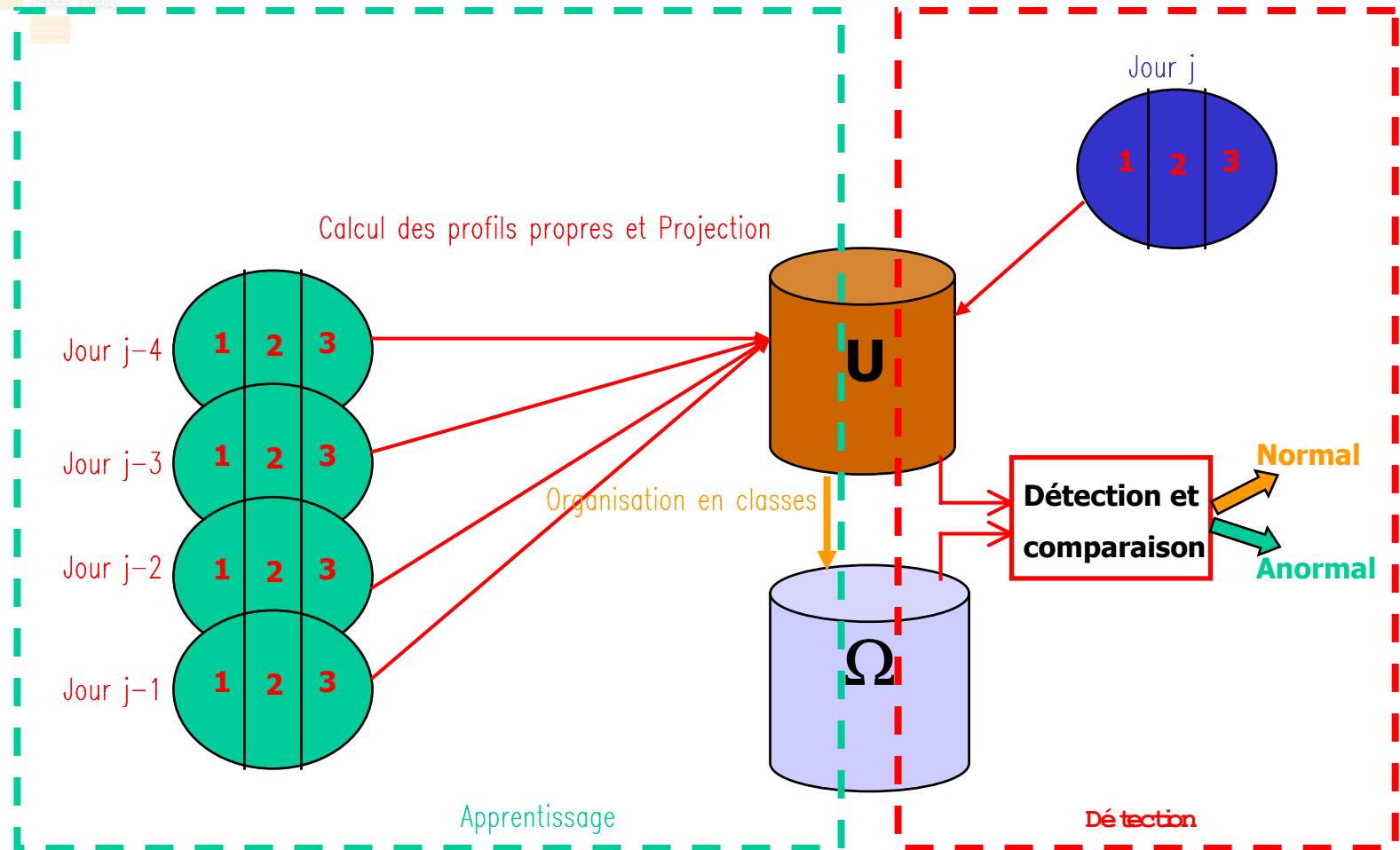
## ■ Premiers Résultats

	<b>Ensemble appris</b>	<b>Ensemble non appris</b>
<b>Identification avec succès</b>	<b>17/24 (70,83%)</b>	<b>10/24 (41,67%)</b>
<b>Mauvais rejet Fausse alerte</b>	<b>7/24 (29,17%)</b>	<b>14/24 (58,33%)</b>

**Peut mieux faire ☹️**

# Application

## Amélioration Proposée: apprentissage glissant



# ■ Application

## ■ Amélioration des Résultats avec l'Apprentissage Glissant

	Ensemble Familier		Ensemble non Familier
	Profils Appris	Profils non Appris	
<i>Taux d'identification avec succès</i>	100% (46/46)	94,44% (34/36)	-
<i>Taux de Confusion ou faux négatifs</i>	0%	0%	0%
<i>Taux de rejet avec succès</i>	-	-	100%
<i>Taux de faux rejet ou faux positifs</i>	0	5,56% (2/36)	-

# ■ Conclusion et Perspectives

---

- Résultats préliminaires encourageants.
  - Méthode simple à implémenter.
  - Apprentissage glissant.
- Travaux futurs
  - Passer à une plus grande échelle (plusieurs dizaines d'utilisateurs)
  - Recherche et application à d'autres nouvelles sources d'audit.

# ■ Des Questions ?

---

