XMeta

Une approche bayésienne pour le computer forensic

Thomas DUVAL thomas.duval@supelec.fr

Bernard JOUGA bernard.jouga@supelec.fr

Laurent ROGER laurent.roger@dga.defense.gouv.fr

Journées Recherche

Supélec

17 mars 2005

Agenda:

- **Objectifs**
- **Réseaux Bayésiens**
  - Définition
  - Exemple
- **Le système XMeta**
  - Architecture
  - Plateforme
- **L'affaire Mitnick**
  - Description
  - Les résultats de XMeta
  - Les résultats de Shimomura
- **Conclusion**

Introduction:

Le faits :

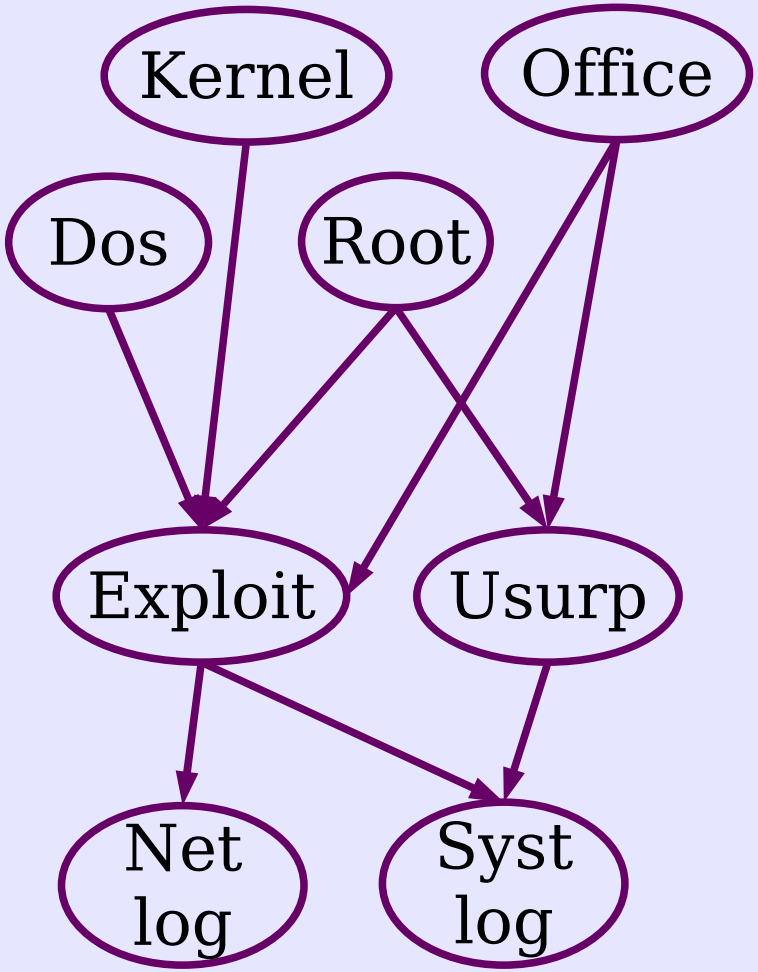
- L'analyse forensique est basée sur la connaissance des experts
- Peu de définitions formelles existent à ce jour [dfrws.org]

Objectifs:

- **CACF** (*Computer-Aided Computer Forensics*)
  - Présentation des logiciels les plus vulnérables
  - Présentation des **attaques** (exploit, usurpation, etc.)
  - Suggestion de possibles **actions** complémentaires (suppression de données, message, etc.)
  - Liste des **techniques d'investigation**
  - Connaissance du **scénario** de l'attaque

- Par un système expert basé sur une approche bayésienne

Réseaux Bayésiens : definition (1/2)



Graphes acycliques orientés :

- les noeuds sont des variables
- les liens sont des relations causales entre les noeuds, pondérés par des probabilités

Lois de probabilité :

- Théorème de Bayes :

```
p(Usurp|Root) = p(Usurp,Root)/p(Root)
```

Construction :

- par la connaissance des experts
- par apprentissage de cas

Algorithme d'inference :

- exact *(Lauritzen-Spiegelhalter)*
- *approximatif (Likelihood weighting)*

Utilisation :

- descendante :

  `Exploit et Usurp ➔ Syst. Log`

- montante :

  Usurp ➔ `Root et/ou Office`

Réseaux Bayésiens : définition (2/2)

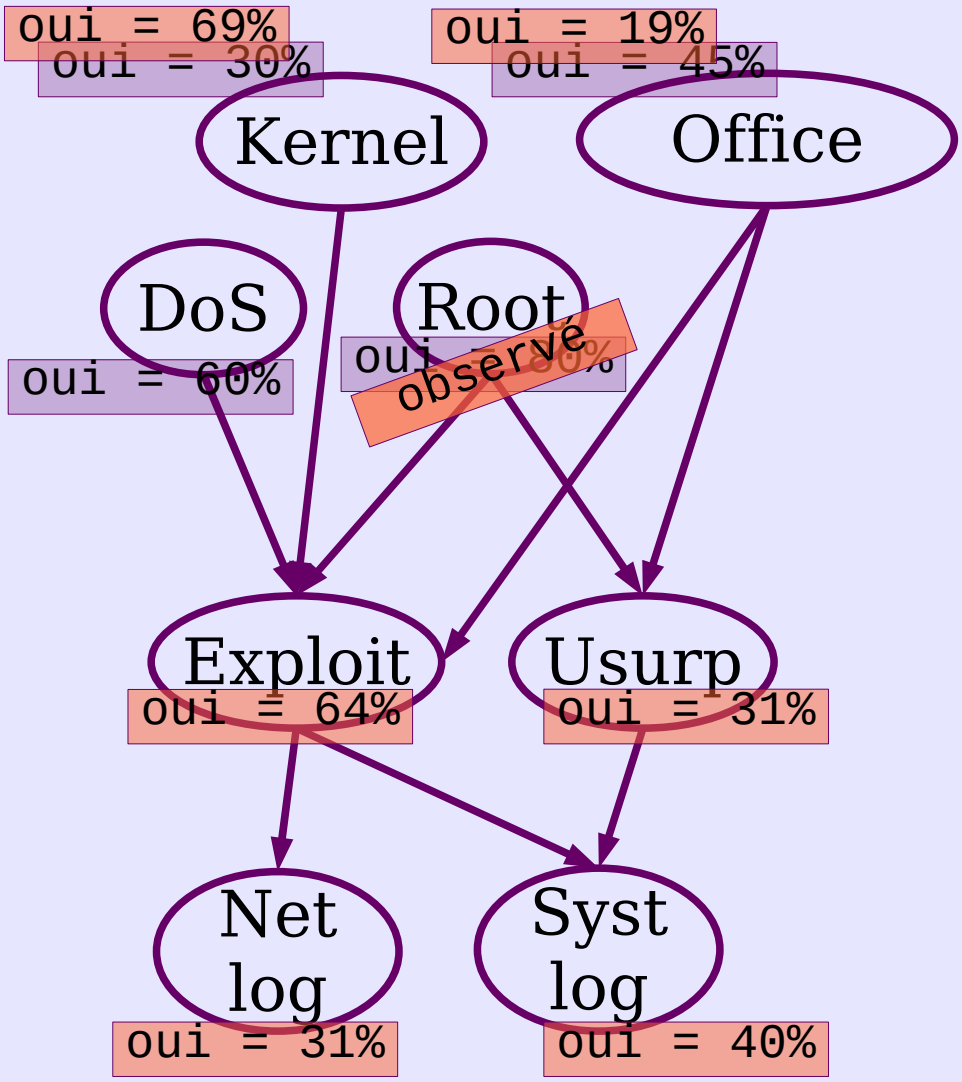
Pourquoi :

- Modélisation de situations complexes
- Logique probabiliste
- Approche robuste quand des données sont manquantes

Exemples d'applications forensiques :

- Analyse de systèmes pour retrouver des données permettant d'identifier des communications suspectes [Costa, Jones, Liao, Malgari]
- Modélisation de crimes (non-informatiques) en proposant des indices sur la culpabilité des personnes impliquées [Levitt, Laskey]

Réseaux Bayésiens : un exemple



Si on observe l'obtention des droits root sur un système Linux :

- mettre le noeud Root à **observé**
- calculer les probabilités
- vérifier les logiciels
- vérifier les attaques
- vérifier les techniques d'investigation

Dans ce cas le résultats est :

- le logiciel attaqué : Kernel
- L'attaque : Exploit
- la meilleure technique d'investigation est de vérifier les logs système

Probabilités initiales : analyse statistique des cas précédents

Probabilités mises à jour : calculé par un algorithme Bayésien

Agenda:

- **Objectifs**
- **Réseaux Bayésiens**
  - Définition
  - Exemple
- **Le système XMeta**
  - Architecture
  - Plateforme
- **L'affaire Mitnick**
  - Description
  - Les résultats de XMeta
  - Les résultats de Shimomura
- **Conclusion**

Le système Xmeta : définitions

On modélise 6 types de noeuds :

- **TH** : *Targeted Hardwares* et **TS** : *Targeted Softwares*
  nomemclature ICAT (plus de 7000 logiciels et versions rescensés)
- **RD** : *Reported Damages*
  nomemclature ICAT

```
LT_Security_protection   LT_Obtain_all_priv   LT_Obtain_some_priv
LT_Confidentiality       LT_Integrity         LT_Avaibility
LT_Sec_prot_other
```

- **GA** : *Generic Attacks*
  nomemclature DGA

```
exploit     broadcast       diversion    usurp        chaff    intercept
listing     embezzlement    bounce       net_listen   listen   troyen
decrypt     parasit         repeat       bypass       degrad   blocking
overrun     brut_force      control
```

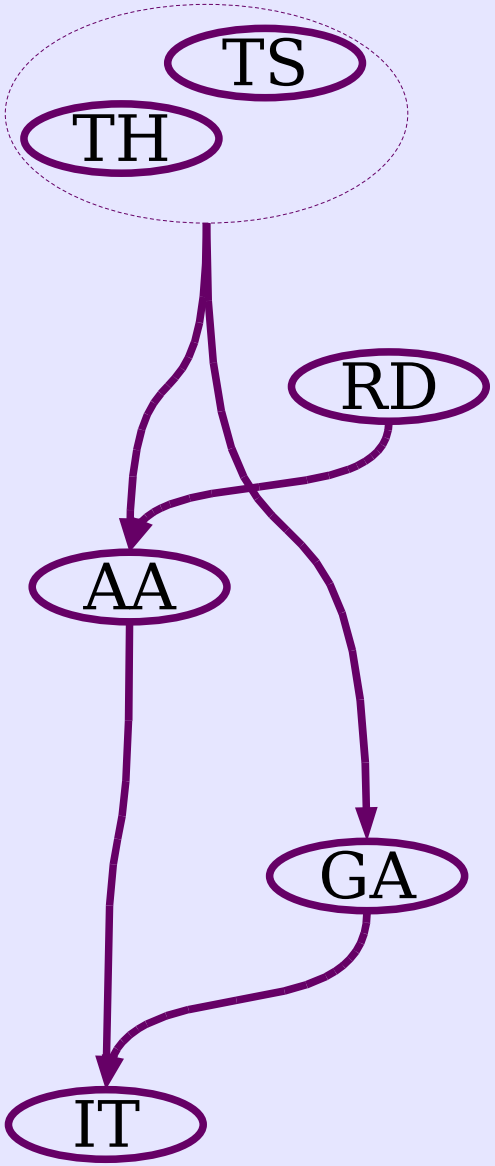
- **AA** : *Additional Attacks*
  nomemclature DGA

```
cnx_illic       attribute    inhib_detect     scan_use
msg             encrypt      hidden_channel   infection
```

- **IT** : *Investigation Techniques*
  nomemclature personnelle

```
image       retrieve     comm     check_syst   log_net      physic
var_syst    topo_int     check_net    topo_ext
```

Le système Xmeta : plan d'investigation

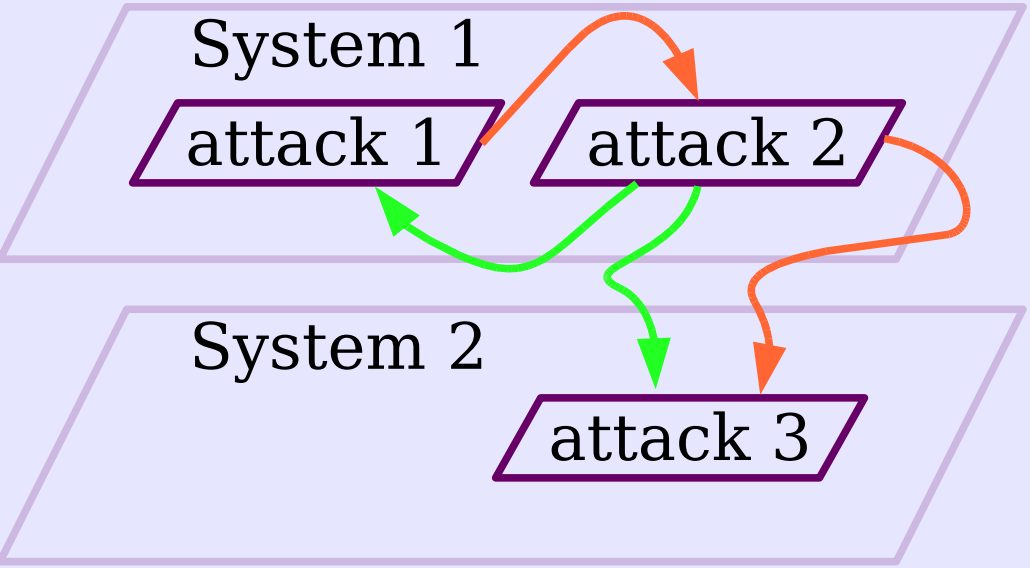


On définit un IP (*Investigation Plan*) pour une configuration système et une attaque données :

- construit à la demande à partir de la configuration du système
- structure extraite de la base ICAT
- tables de probabilité créées avec l'algorithme **K2**
- probabilités postérieures calculées par l'algorithme d'inférence approximative **LW** (`Likelihood Weighting`)

Le système Xmeta : Liens entre IP

Les liens matérialisent la progression de l'attaque :



point de vue de l'attaquant

point de vue de l'enquêteur

Le système Xmeta : plateforme

Moteur d'inférence :

- Version 0 : BNJ (*Bayesian Network Tools in Java*) (bndev.sourceforge.net)
- Version 1 : implémentation spécifique “Python BN“

IHM :

- Python + GTK

Sauvegarde des données :

Un format spécifique a été développé :

- **CFXR** (*Computer Forensics XML Report*)

Expérimentations:

Les premières ont été faites avec des cas triviaux :

- bons résultats
- mais besoin de cas réels plus complexes

Agenda:

- **Objectifs**
- **Réseaux Bayésiens**
  - Définition
  - Exemple
- **Le système XMeta**
  - Architecture
  - Plateforme
- **L'affaire Mitnick**
  - Description
  - Les résultats de XMeta
  - Les résultats de Shimomura
- **Conclusion**

L'affaire Mitnick : introduction

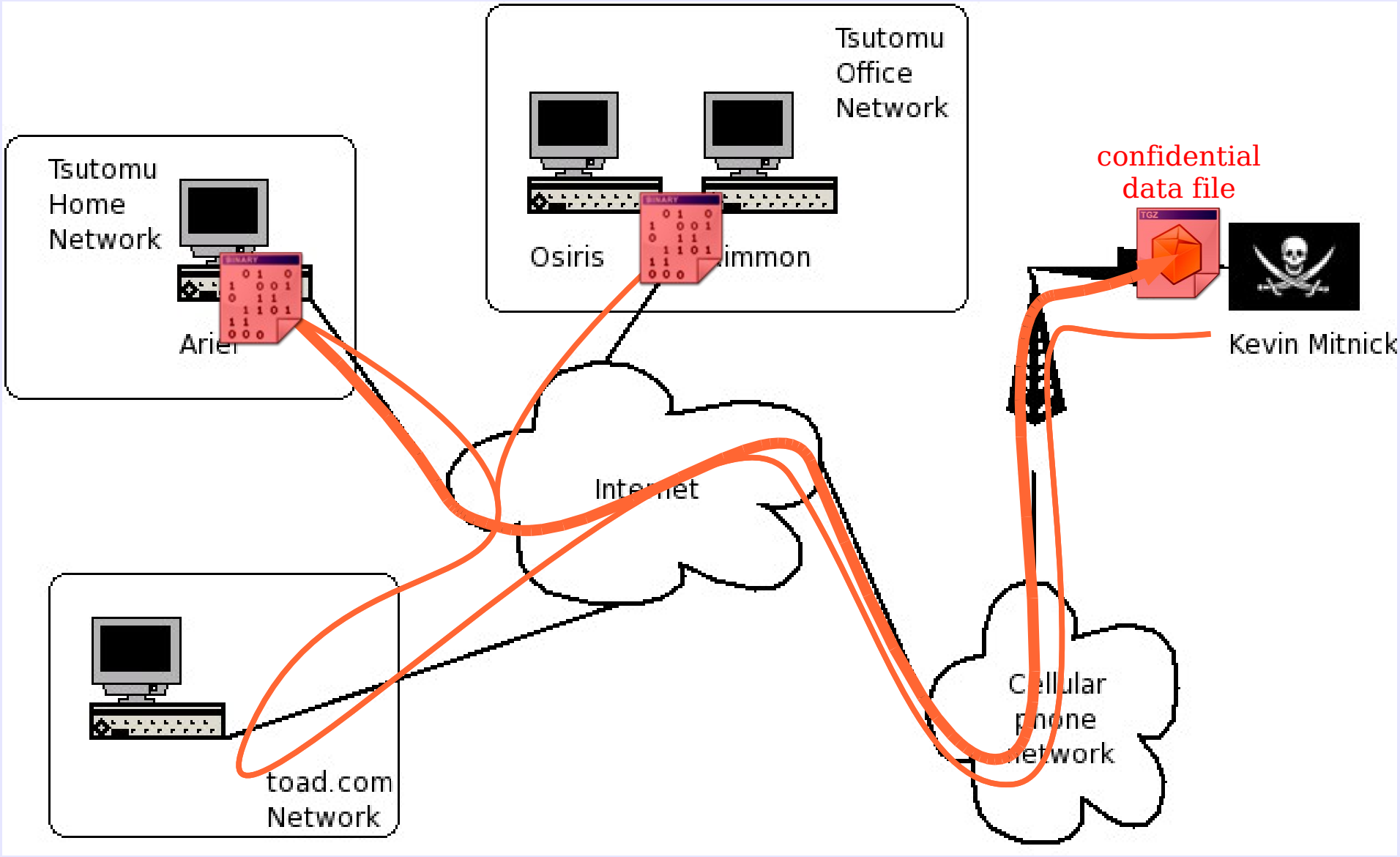
Les équipements de Tsutomu Shimomura ont été attaqués par un inconnu en décembre 1994.

Tsutomu Shimomura est *senior fellow* au Supercomputer Center de San Diego.

Après 7 semaines de traque, Kevin Mitnick a été arrêté puis accusé d'intrusions informatiques et de vol de données.

L'attaque de Mitnick est intéressante car elle est complexe et de nombreux systèmes informatiques sont impliqués.

L'affaire Mitnick : description



L'affaire Mitnick : XMeta (1/3)

Configuration:
SunOS
GNU Tar
GNU Ghostscript
fingerd
ruserd
FTP

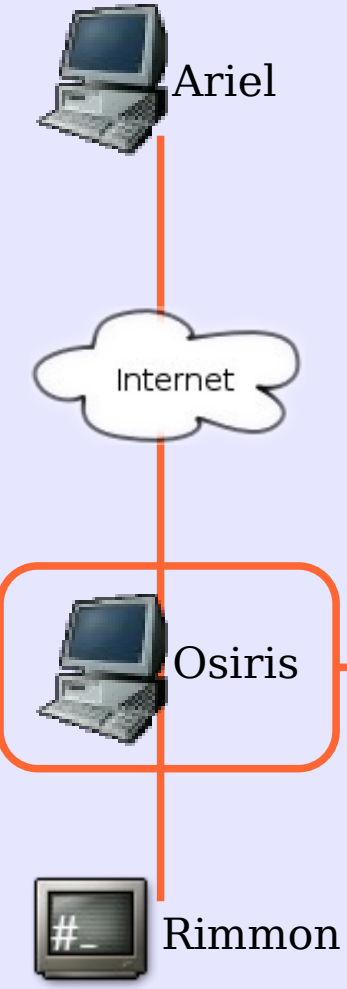
Observation : un fichier a été téléchargé depuis Ariel
- LT_Confidentiality

Attacks:
- bypass a security element (65%)
- diversion (56%)
- brut_force (56%)

Softwares:
- GNU tar (73%)
- Finger Service (73%)
- FTP (27%)

- L'attaque vient certainement de l'extérieur car les attaques ci-dessous ne peuvent pas être locales.

L'affaire Mitnick : XMeta (2/3)



Configuration:
SunOS
GNU Tar
GNU Ghostscript
fingerd
ruserd
FTP

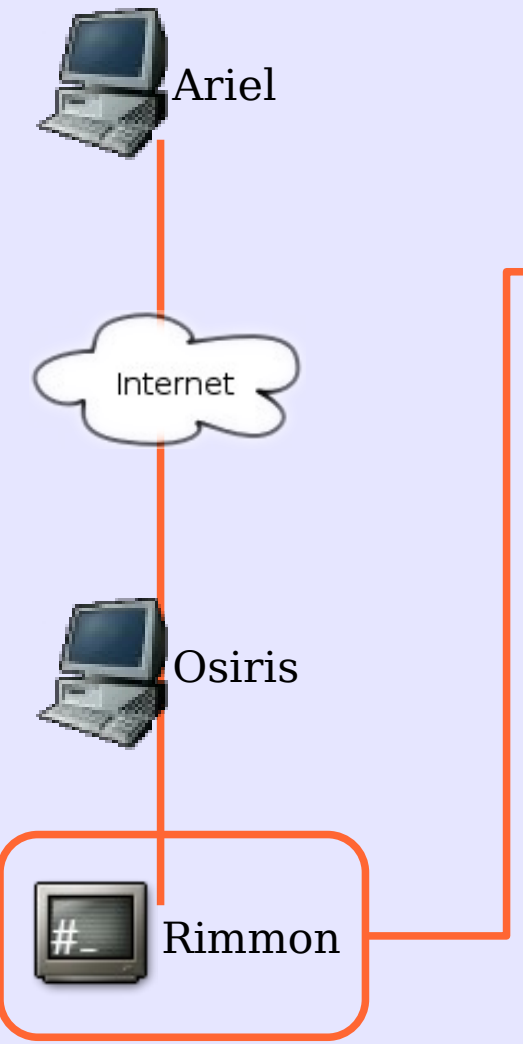
Observation : étrange comportement d'Osiris : il est déconnecté d'Ariel
- LT_Availability

Attacks:
- repeat (100%)
- overrun (89%)
- bypass (68%)

Softwares:
- FTP (73%)
- GNU Tar (38%)
- GNU Ghostscript (38%)

- Comme pour Ariel, l'attaque ne peut être locale.

L'affaire Mitnick : XMeta (3/3)



Configuration:
SunOS
GNU Tar
GNU Ghostscript
fingerd
ruserd
FTP

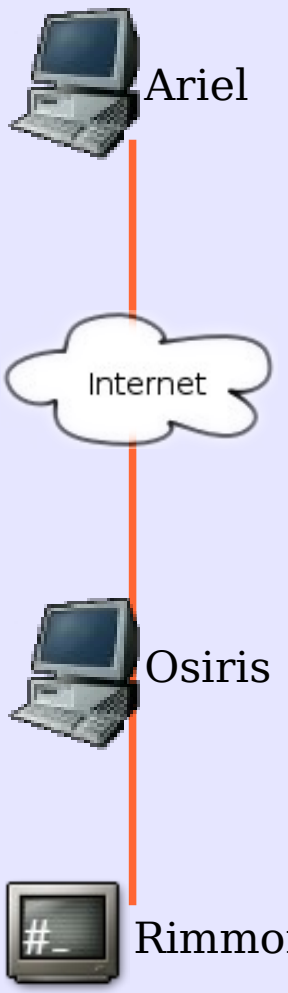
Observation :
Shimomura découvre que l'attaquant a installé un module dans le système :
- LT_Obtain_all_priv
- LT_Availability

Attacks:
- trojan (93%)
- bypass (78%)
- brut_force (58%)
- ...
- 10/21 **overrun**

Softwares:
- FTP (59%)
- GNU Tar (41%)

- On a une attaque en 3 phases, la prochaine étape serait d'enquêter sur la date des attaques.

L'affaire Mitnick : Synthèse

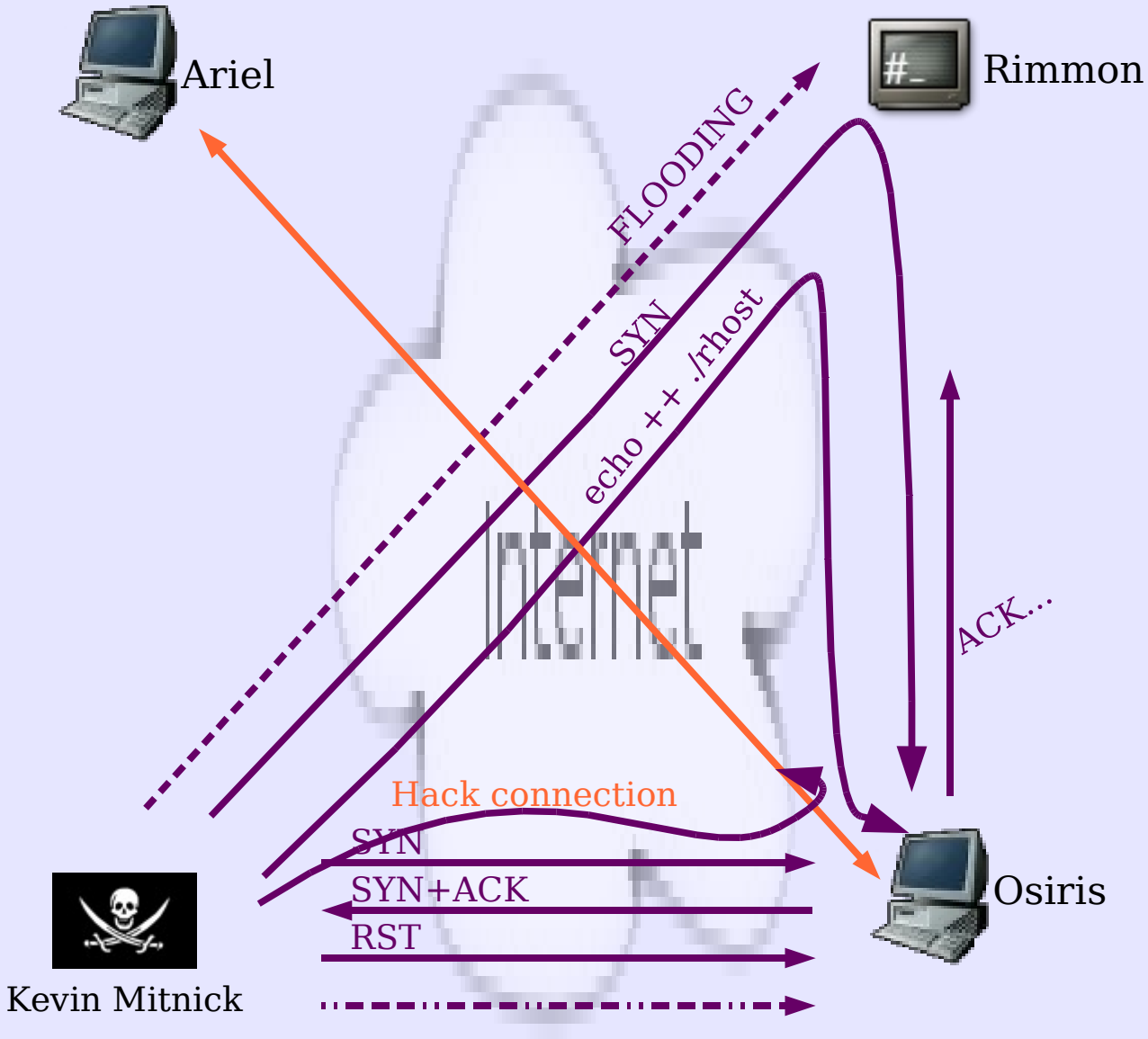


- un fichier a été téléchargé depuis Ariel grâce à une attaque de type **bypass** ou diversion
- l'attaquant a certainement utilisé l'attaque **repeat** sur Osiris et / ou **bypass** pour entrer
- L'attaquant a utilisé les relations de confiance entre Osiris and Rimmon pour accéder à Osiris

Les questions qui restent en suspens :

- Qui a pénétré dans le réseau de toad.com ?
- Comment l'attaquant a obtenu un accès root sur Osiris ?
- Quelle est la destination finale du fichier volé ?

L'affaire Mitnick : Les résultats de Shimomura



1) Récupération des numéros de séquences valides d'Osiris

2) inondation de Rimmon et envoie d'une commande à Osiris

3) installation du module dans la mémoire d'Osiris pour récupérer la connexion avec Ariel

Conclusion

Xmeta est un système expert utilisant les réseaux Bayésiens
Il peut :

- donner des indices à propos des :
  - attaques
  - actions complémentaires
  - logiciels
- proposer les meilleures techniques d'investigations à utiliser
- être mis en oeuvre sur des cas non triviaux

http://www.rennes.supelec.fr/ren/perso/tduval/

Travaux à venir

Amélioration de la base de données (techniques d'investigation)

Expérimentations sur des cas réels (avec la DGA)

Travail sur les liens entre plans pour :

- décrire des attaques multiples
- prédire la prochaine ou la précédente étape d'une attaque

Travail sur les profils d'attaquant :

- caractériser les attaquants (*insiders* / *outsiders*)
- reconnaître des attaquants typiques (*script kiddies*, expert, etc.)

http://www.rennes.supelec.fr/ren/perso/tduval/

Publications

Thomas Duval, Bernard Jouga, Laurent Roger. XMeta: a Bayesian approach for computer forensics. Work in Progress Session of the Annual Computer Security Applications Conference (ACSAC). Tucson (USA). December 2004.

Thomas Duval, Bernard Jouga and Laurent Roger. The Mitnick Case: how Bayes could have help. First Annual IFIP WG 11.9 International Conference on Digital Forensics. Orlando (USA). February 2005.

Thomas Duval, Bernard Jouga et Laurent Roger. XMETA : une approche bayésienne pour le computer forensics. Symposium sur la Sécurité des Technologies de l'Information et des Communications. Rennes. Juin 2005.

XMeta system : prototype