Réseaux bayésiens pour le filtrage d'alarmes dans les systèmes de détection d'intrusions

Ahmad Faour[1,2] ahmad.faour@ul.edu.lb
Philippe Leray[1] philippe.leray@insa-rouen.fr
Cédric Foll[1,3] cedric.foll@ac-rouen.fr

[1] Laboratoire PSI - FRE CNRS 2645, INSA Rouen, France
[2] Laboratoire LPM, Université Libanaise, Beyrouth, Liban
[3] Rectorat de Rouen, France

1 Introduction

La détection des tentatives d'attaques sur un réseau est une problématique très importante dans le domaine de la sécurité informatique. Les NIDS (*Network Intrusion Detection Systems*), systèmes de détection d'intrusions, génèrent tellement d'alertes sur un réseau qu'il en devient très difficile de déterminer celles générées par une attaque réelle. L'utilisation d'outils de raisonnement probabiliste comme les réseaux bayésiens (RB) peut être efficace pour détecter les problèmes réels. Nous allons donc tout d'abord présenter les systèmes de détection d'intrusions et leurs limites puis passer brièvement en revue l'application de méthodes d'apprentissage à cette problématique. Nous décrirons enfin notre architecture de filtrage d'alarmes issues de NIDS.

2 Systèmes de détection d'intrusions

Les firewalls utilisés sur les réseaux TCP/IP fonctionnent sur l'analyse des couches IP et TCP/UDP/ICMP, pour déterminer quelles sont les machines impliquées dans la connexion et à quel service la connexion s'adresse. Ce genre d'approche, bien que nécessaire, se révèle insuffisant dans bien des cas (Chambet, 2002). Il faut donc pousser plus loin l'analyse en examinant aussi les couches réseaux supérieures. Cette tâche, plus difficile, est dévolue aux NIDS. Ces logiciels fonctionnent le plus souvent par signatures, sur le même principe que les anti-virus (Zimmermann et al., 2002), en répertoriant les attaques connues. Une alarme est donc générée à chaque fois qu'une trame réseau ressemble à une des attaques répertoriées. Lorsqu'un nouvel exploit (tentative d'intrusion réussie) est répertorié, une signature adaptée sera ajoutée à la base de signatures. Cette approche est souvent utilisée conjointement avec une approche statistique dans laquelle le NIDS détermine d'abord un profil type du réseau (nombre de paquets échangés, volume des flux, nombre de connections, etc.) et alarme ensuite l'administrateur lorsque le trafic courant dévie de ce profil. Malheureusement, les NIDS émettent généralement une quantité importante d'alarmes que l'administrateur n'est pas capable d'interpréter rapidement.

Depuis (Denning, 1987), les approches à base d'apprentissage statistique proposées pour la détection d'intrusion peuvent être classées en deux types : les méthodes essayant d'opérer avec les mêmes informations que les NIDS classiques (analyse de données réseaux), et celles opérant à partir de données comportementales de plus haut niveau (fichiers de logs de certaines applications ou du système).