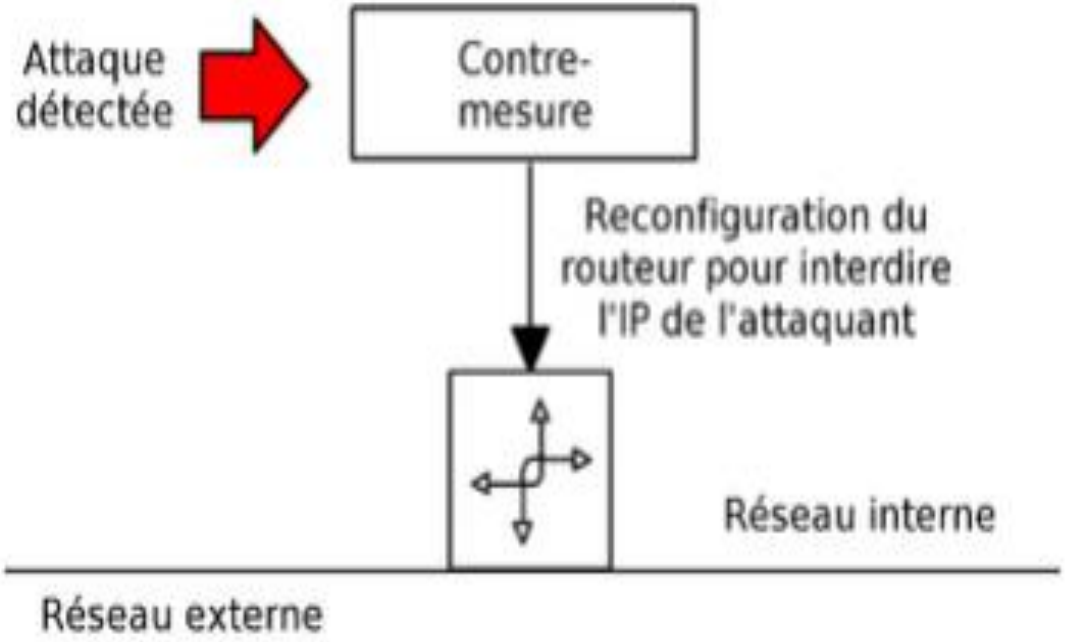
Les systèmes de détection d’intrusion (IDS)

Quelques définitions

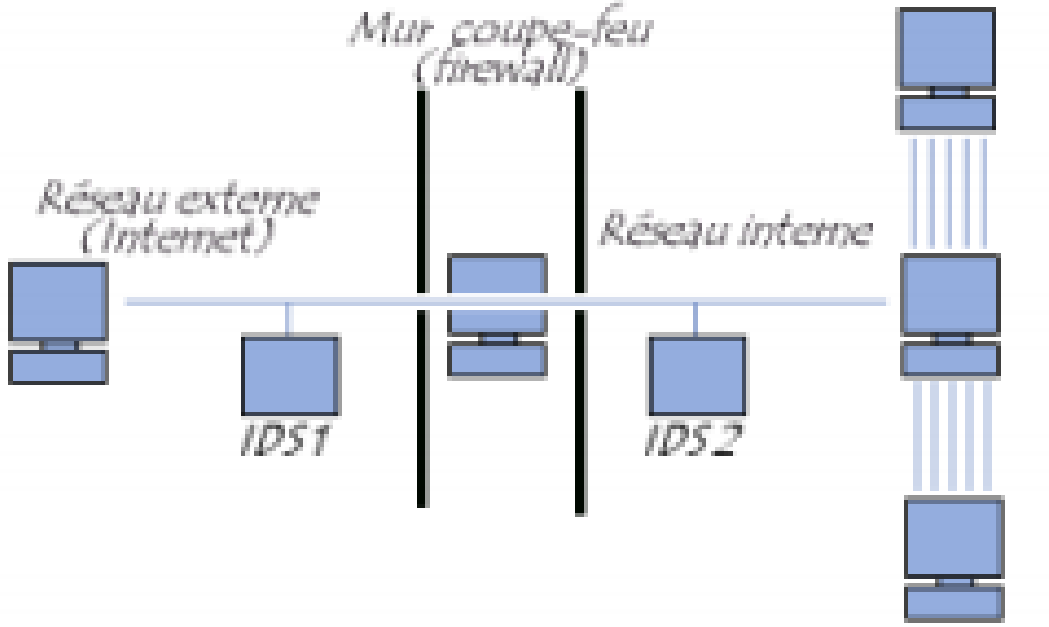
- **Intrusion** : C'est une violation d'une politique de sécurité d'un système donnée, c'est-à-dire une violation d'une des propriétés de confidentialité, d'intégrité ou de disponibilité du système en question.
- **Attaque** : A ne pas confondre avec une intrusion. Une attaque est une tentative de violer la politique de sécurité alors qu'une intrusion est une violation effective de cette politique. En d'autres termes, une intrusion est une attaque réussie.
- **IDS**: C'est un mécanisme écoutant le trafic réseau de manière furtive afin de repérer des activités anormales ou suspectes. La détection d'une attaque se traduit par le déclenchement d'une alarme sur le réseau.
- **IPS**: Les IPS sont des IDS actifs : En cas de détection d'une attaque, l'IPS ne se contente pas de notifier l'administrateur réseau mais agit pour bloquer ou corriger les risques d'intrusion. Une action de prévention peut être, par exemple, le blocage de l'adresse IP du présumé attaquant.



http ://www.labo-cisco.com/index.php ?option=com_content&task=view&id=951&Itemid=11

Classification des IDS

- Les méthodes pour détecter les intrusions reposent sur l'observation d'un certain nombre d'événements et l'analyse de ceux-ci.
- Il s'agit premièrement de collecter les informations que l'on souhaite analyser.
- Ces informations proviennent des fichiers de journalisation du système ou des sondes mises en place par les outils de détection d'intrusions tels les « sniffers » réseau.
- On peut classer les IDS en trois grandes catégories : les IDS réseaux (network-based DS ou NIDS), les IDS systèmes (host-based IDS ou HIDS) et les IDS dits «hybrides ».

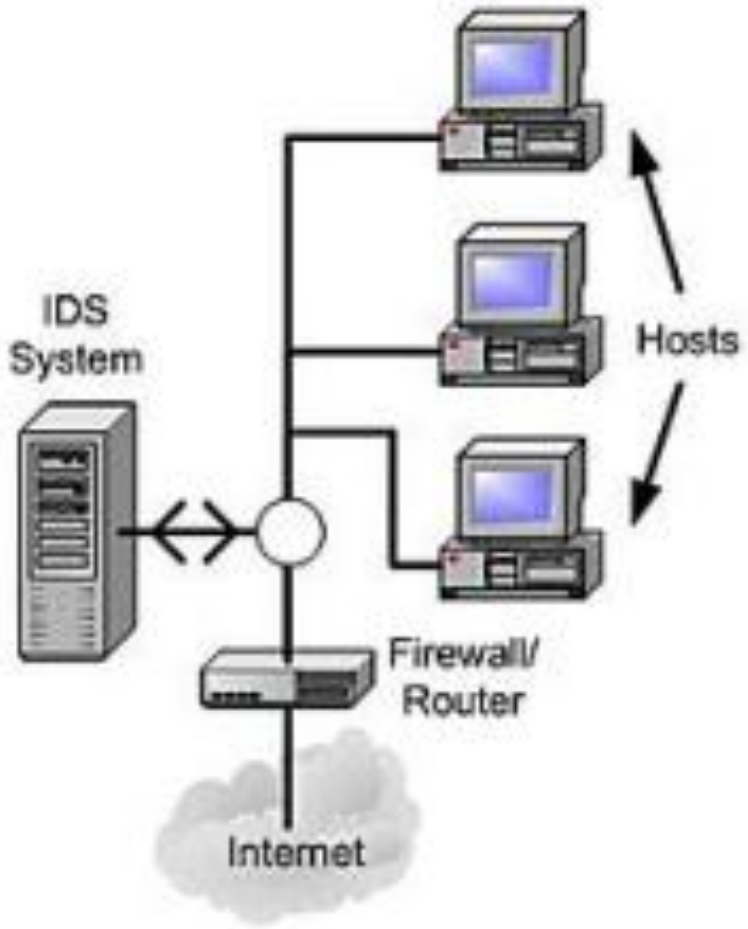


http ://www.commentcamarche.net/contents/detection/ids.php3

Déploiement de plusieurs IDS

1- IDS Réseaux (ou NIDS) :

- Les IDS réseaux (Network IDS) analysent en temps réel le trafic qu'ils aspirent à l'aide d'une sonde (carte réseau en mode "promiscuous"; Promiscuous mode (ou mode promiscuité) se réfère à une configuration de la carte réseau qui permet à celle-ci d'accepter tous les paquets qu'elle reçoit, même si ceux-ci ne lui sont pas adressés). Ensuite, les paquets sont décortiqués puis analysés.
- Il est fréquent de trouver plusieurs IDS sur les différentes parties du réseau. On trouve souvent une architecture composée d'une sonde placée à l'extérieur du réseau afin d'étudier les tentatives d'attaques et d'une sonde en interne pour analyser les requêtes ayant traversé le pare-feu.

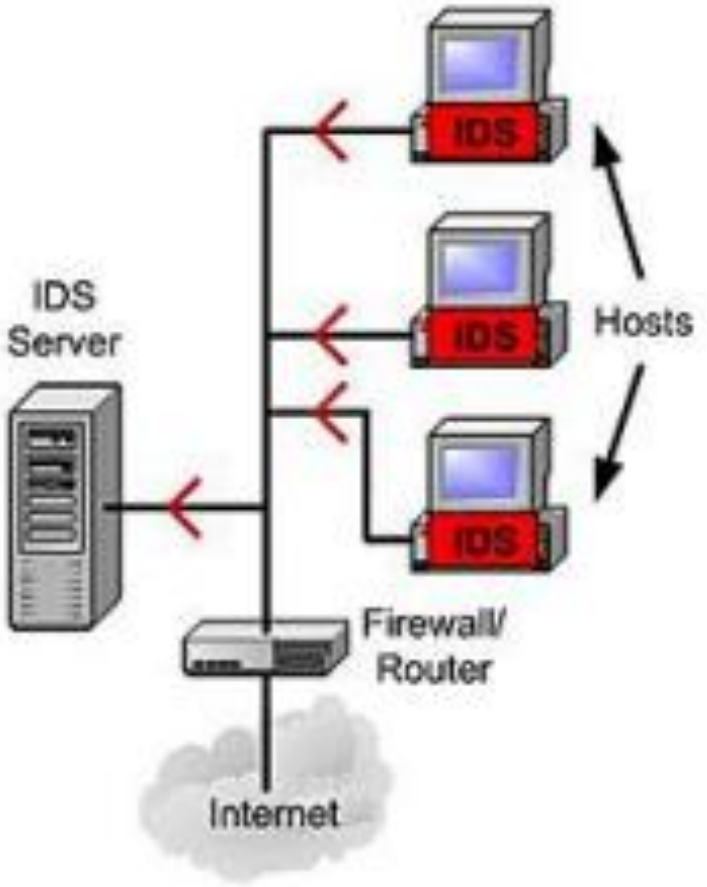


http ://www.informit.com/articles/article.aspx ?p=29601

Architecture d'un NIDS

2- IDS Systèmes (ou HIDS)

- Les IDS systèmes (Host IDS) analysent le fonctionnement et l'état des machines sur lesquelles ils sont installés afin de détecter les attaques en se basant sur des démons (Un démon est un type particulier de programme informatique, un processus qui s'exécute en arrière-plan plutôt que sous le contrôle direct d'un utilisateur; tels que syslogd par exemple).
- L'intégrité des systèmes est alors vérifiée périodiquement et des alertes peuvent êtres levées.

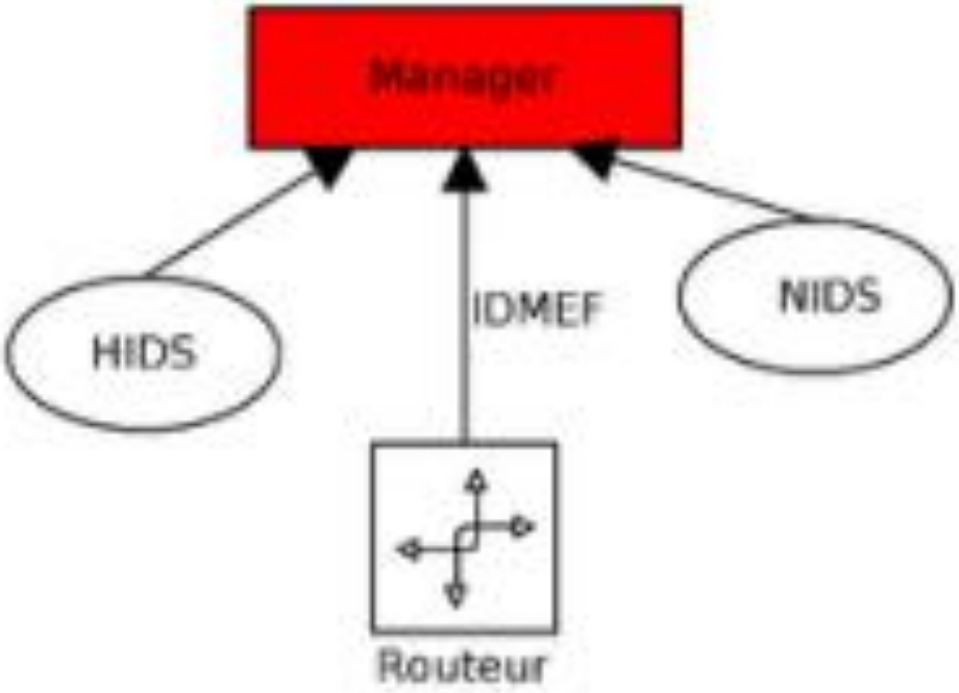


http ://www.informit.com/articles/article.aspx ?p=29601

Architecture d'un HIDS

3- IDS Hybrides

- Les IDS hybrides rassemblent les caractéristiques des NIDS et HIDS. Ils permettent, en un seul outil, de surveiller le réseau et les terminaux. Les sondes sont placées en des points stratégiques, et agissent comme NIDS et/ou HIDS suivant leurs emplacements.
- Toutes ces sondes remontent alors les alertes à une machine qui va centraliser le tout, et agréger/lier les informations d'origines multiples.
- Ainsi, on comprend que les IDS hybrides sont basés sur une architecture distribuée, où chaque composant unifie son format d'envoi (par exemple IDMEF). Cela permet de communiquer et d'extraire des alertes plus pertinentes.
- Les avantages des IDS hybrides sont multiples :
  - Moins de faux positifs
  - Meilleure corrélation (la corrélation permet de générer de nouvelles alertes à partir de celles existantes)
  - Possibilité de réaction sur les analyseurs



http ://doc.ubuntu-fr.org/utilisateurs/psychederic/pare-feu

Principe de l'IDS hybride

Méthodologie de détection

- Plusieurs approches de détection d'intrusion sont utilisées par l'IDS (conjointement ou exclusivement). Les méthodologies qui vont être traitées dans ce cours sont les méthodes basées signatures et les méthodes comportementales.

1. Les IDS à signatures (ou à scénarios)

1.1 **Principe**

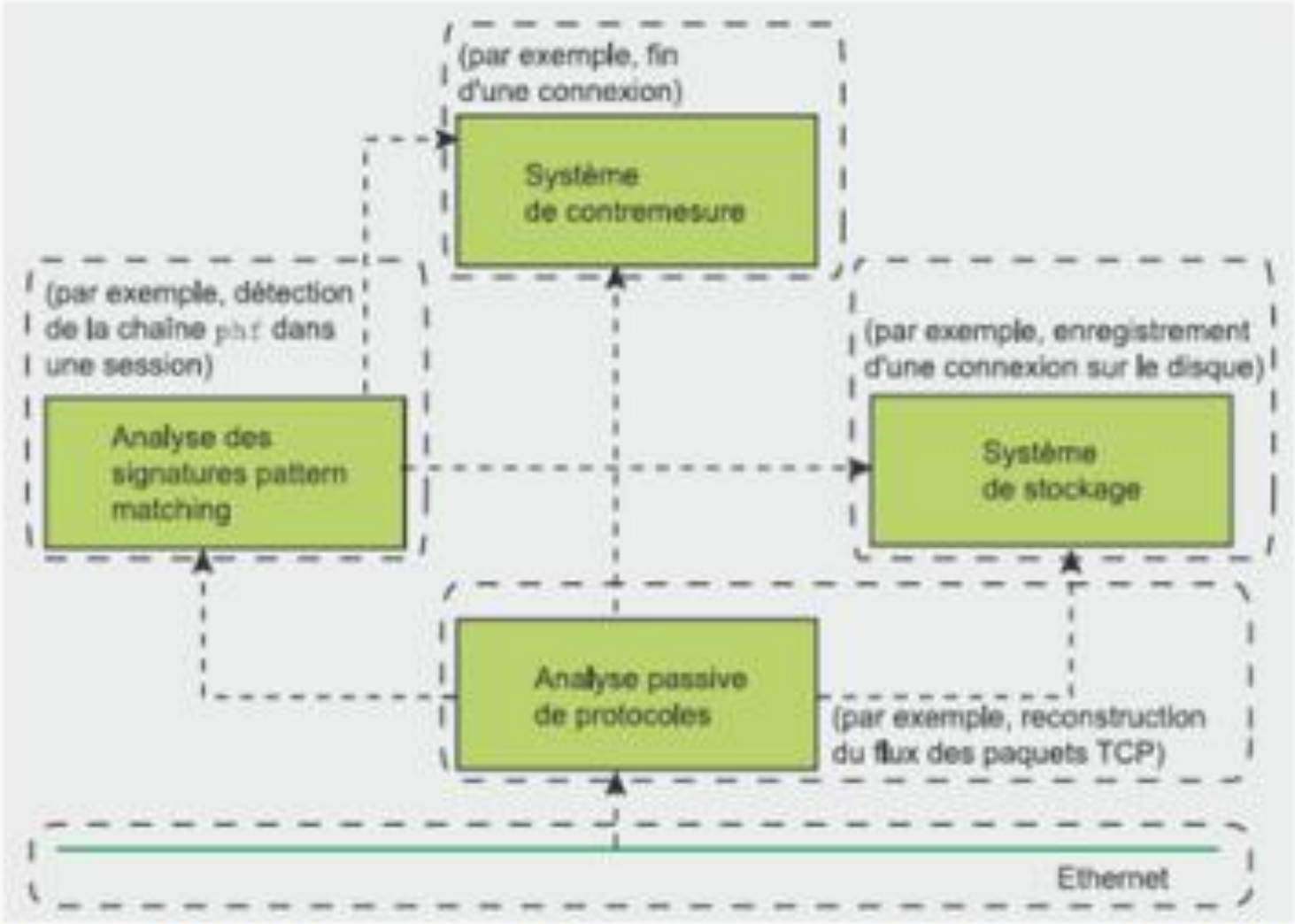
- Une signature représente le scénario d'une attaque.
- Cette approche consiste à rechercher dans l'activité de l'élément surveillé (un flux réseau) les empreintes d'attaques connues, à l'instar des anti-virus.
- Trois familles de méthodes sont utilisées par les IDS à signature qui se basent tous sur la recherche d'un profil connu d'attaque.

1.2 Systèmes Experts

- Un système expert est un système basé sur trois types de règles.
- Le premier type sert à coder ce qui est suspect à priori (par rapport à la politique de sécurité mise en œuvre).
- Un deuxième type qui concerne les failles et les vulnérabilités connus d'un système et qui sont, en général, publiés par des organismes internationaux (comme le CERT ou Computer Emergency Response Team).
- Le dernier type est utilisé pour coder le savoir faire de l'administrateur réseau.

1.3 “Pattern matching” (reconnaissance de formes)

- Cette méthode consiste à identifier dans les paquets analysés une suite d’événements ou de caractères caractéristiques d’une attaque connue.
- En fait, Le trafic réseau peut être vu comme une chaîne de caractères principale et les scénarios d’attaque comme des sous-suites qu’on veut identifier.
- La figure ci-dessous montre l’architecture générale d’un IDS à «pattern matching » :



"*Les systèmes de détection d'intrusion vus de l'intérieur*", *Antonio Merola*

Architecture d'un IDS utilisant le pattern matching

- Un exemple de chaînes de caractères malicieuses est la chaîne “/scripts/iisadmin/default.htm” qui vise à accéder à la page d’administration d’un serveur Web IIS (Internet Information Services).
- Un exemple d’une règle Snort (SNORT est un IDS réseau gratuit disponible à l’adresse suivante : http ://www.snort.org/) pour contrer cette menace est donné ci-dessous :

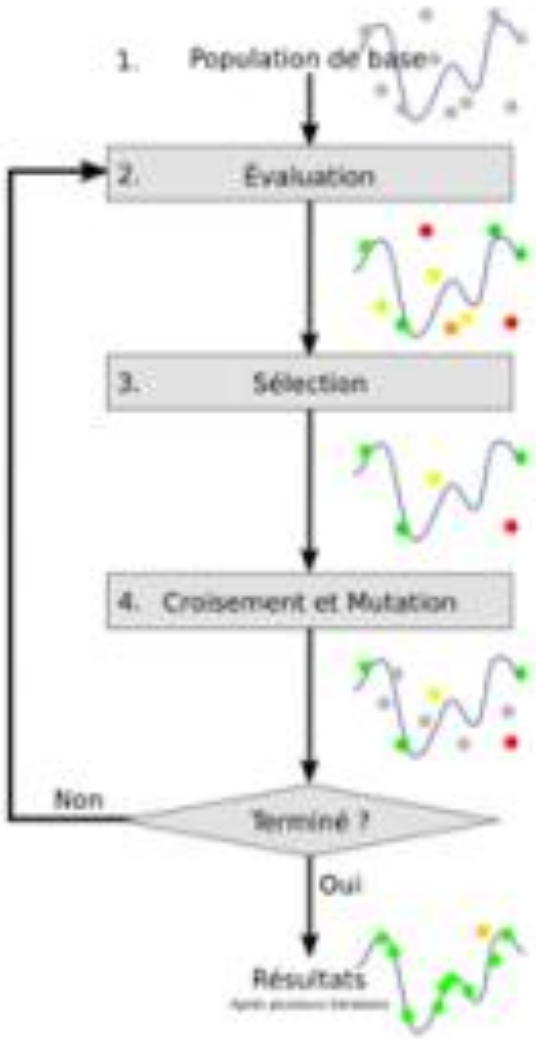
```
alert tcp $EXTERNAL_NET any -> $HTTP_SERVERS $HTTP_PORTS (msg:"WEB-IIS /scripts/
iisadmin/default.htm access"; flow:to_server,established; uricontent:"/scripts/i
isadmin/default.htm"; nocase; classtype:web-application-attack; sid:994;  rev:6;
)
```

Une règle Snort...

1.4 Algorithmes génétique

- Les algorithmes génétiques utilisent la notion de sélection naturelle et l'appliquent à une population de solutions potentielles à un problème difficile (dont on ne sait pas trouver la solution optimale) pour trouver une solution approchée dans un temps raisonnable.
- Dans notre cas la population de départ peut être des règles de détection basiques. A travers l'algorithme génétique, on génère d'autres règles qui couvrent au mieux les cas des flux anormaux.

- Basé sur les logs récoltés par le sniffeur de réseaux, la première étape est de récupérer des séries de données contenant toutes les informations nécessaires pour créer des règles.
- Celles-ci peuvent inclure l'adresse IP source, l'adresse IP destination, le port source, le port destination, le protocole utilisé ainsi qu'un champ précisant si cette connexion peut être considérée comme une intrusion ou non.
- La première partie de l'algorithme génétique va en fait agir comme un algorithme de recherche servant à faire correspondre les connexions à des règles déjà établies.
- Ensuite, la seconde étape est de déterminer si une règle est «bonne» ou « mauvaise ».
- Ceci est réalisé à l'aide de valeurs que l'on attribue à chaque règle et qui sont recalculées lors d'une détection d'intrusion.
- Enfin, les mauvaises règles sont remplacées par des nouvelles règles issues de fusions de « bonnes » règles.



http ://fr.academic.ru/dic.nsf/frwiki/81706

Principe de l'algorithme génétique

1.5 Discussion

Les IDS à signature présentent les principaux avantages suivants :

- Déclencher des alertes pertinentes qui rendent compte du type attaque voire de la vulnérabilité exploitée. Ceci est particulièrement important pour l'administrateur réseau qui, en se basant sur la base de connaissances, peut entamer les actions adéquates pour restaurer l'état initial du système.
- Les IDS à signature sont en général très performants de point de vue utilisation des ressources surtout quand ils sont des NIDS. l'IDS ne fonctionne pas sur les machines rendant un service dans le système. Les performances des services rendus par le système ne sont pas affectées significativement par la présence de l'IDS.
- Une relative simplicité de mise en œuvre de cette technologie. Des logiciels assez performants sont disponibles en open source ce qui fait qu'on trouve une grande quantité de documentation sur le sujet.

- Malgré les avantages évoqués, les IDS à signature souffrent d'un gros désavantage qui est la non possibilité de détecter une intrusion non connue même si elle est basique !
- Ces IDS doivent toujours être maintenus et mis à jour pour une protection optimale ce qui peut être délicat.

2. Les IDS comportementaux

- Les IDS comportementaux ont pour principale fonction la détection d'anomalie.
- Leur déploiement nécessite une phase d'apprentissage pendant laquelle l'outil va apprendre le comportement "normal" des flux applicatifs présents sur son réseau.
- Ainsi, chaque flux et son comportement habituel doivent êtres déclarés ; l'IDS se chargera d'émettre une alarme, si un flux anormal est détecté, et ne pourra bien entendu, spécifier la criticité de l'éventuelle attaque.
- Cette technique peut être appliquée non seulement à des utilisateurs mais aussi à des applications et services.
- Plusieurs métriques sont possibles pour déterminer le caractère normal ou non d'un trafic : la charge CPU, le volume de données échangées, le temps de connexion sur des ressources, la répartition statistique des protocoles et applications utilisés, les heures de connexion, etc. . .
- Il existe différentes techniques pour repérer les attaques.

2.1 Approche probabiliste

- On prévoit quelle est la probabilité d'avoir un évènement après un autre.
- Ex : quelqu'un qui se connecte à un site : forte probabilité que la demande de connexion soit suivie de GET http ://www.google.fr HTTP/1.0 Et on peut supposer que ce sera suivi de HTTP/1.1 200 OK Si ce n'est pas ça la plupart du temps, on peut avoir un doute. . .
- Avantages :

  1. Construction du profil simple et dynamique

  2. Réduction de faux positifs
- Inconvénients :

  1. Risque de déformation progressive du profil par des attaques répétées

2.2 Approche statistique

- Cette approche consiste à effectuer des tests sur d'autres éléments concernant l'utilisateur :
  - Le taux d'occupation de la mémoire
  - L'utilisation des processeurs
  - La valeur de la charge réseau
  - Le nombre d'accès à l'Intranet par jour
- Avantages :
  1. Permet de détecter des attaques inconnues
  2. Habitudes des utilisateurs apprises automatiquement
- Inconvénients :
  1. Complexité en termes de maintenance
  2. Beaucoup de faux positifs

2.3 Immunologie

- Cela consiste à établir un modèle de comportement normal des services (et non des utilisateurs).
- L'approche immunologique tente de calquer le comportement du système immunologique pour faire la différence entre ce qui est normal (le soi) et ce qui ne l'est pas (le non-soi).
- Le système immunologique montre en effet beaucoup d'aspects intéressants comme son mode d'opération distribué (il n'y a pas de système de contrôle central) qui lui permet de continuer à fonctionner même après des pertes, sa capacité à apprendre automatiquement de nouvelles attaques pour mieux réagir les prochaines fois qu'elles se présente, sa capacité à détecter des attaques inconnues, etc.
- On peut voir l'approche immunologique de la détection d'anomalies comme une méthode de détection d'anomalie où l'on utilise les techniques de détection des malveillance.

- En effet, les techniques de détection d'anomalie connaissent ce qui est bien et vérifient en permanence que l'activité du système est normale, alors que les techniques de détection de malveillance connaissent ce qui est mal et sont à sa recherche.
- L'approche immunologique propose de rechercher ce qui est mal en connaissant ce qui est bien.
- Avantages :

  1. Nouvelle approche : on ne vérifie pas ce qui est bien, on cherche ce qui est anormal
- Inconvénients :

  1. Nécessite une longue période d'apprentissage

  2. Notion de tiers exclus en logique : on n'obtient pas ce qui est mal en prenant la négation de ce qui est bien

2.4 Discussion

- Les IDS comportementaux offrent une construction du profil de trafic simple et dynamique, aucune configuration n'est réellement requise.
- De plus, les nouvelles attaques sont directement prises en compte (en effet, il n'y a pas de base de signatures).
- C'est une méthode particulièrement adapté dans le cadre d'un trafic est très régulier, la période d'apprentissage est donc favorisée, que cela soit au niveau de la durée, qu'au niveau de la corrélation des protocoles qui entrent en jeu.
- Cependant, ils possèdent quelques désavantages :

- Le bruit peut sévèrement dégrader l'efficacité de ces IDS. Ils sont en effet caractérisés par un fort taux de fausses alarmes : les faux positifs sont les erreurs les plus rependues et cela crée des données parasites qui rendent le système difficile à utiliser voir inutilisable.

  • Le comportement et les activités des utilisateurs pendant la phase d'apprentissage peuvent ne pas être assez statiques pour offrir un comportement normal sur lequel basé une détection d'intrusion, rendant l'IDS inefficace.

  • Un utilisateur peut changer lentement de comportement dans le but d'habituer le système à un comportement intrusif.

  • Enfin, le système peut très bien subir une attaque pendant la phase d'apprentissage.

Les outils disponibles

1. **Critères de choix**

- Aujourd'hui les systèmes de détection d'intrusion sont réellement devenus indispensables.
- Ils s'intègrent donc toujours dans un contexte et une architecture qui imposent des contraintes pouvant être très diverses.
- C'est pourquoi il n'existe pas de grille d'évaluation unique pour ce type d'outil.
- Pourtant un certain nombre de critères peuvent être relevés.

  • **Fiabilité** : Un détecteur d'intrusion doit être fiable ; les alertes qu'il génère doivent être justifiées et aucune intrusion ne doit pouvoir lui échapper.

  • **Réactivité** : Un IDS doit être capable de détecter les nouveaux types d'attaque le plus rapidement possible ; pour cela il doit rester constamment à jour. Des capacités de mise à jour automatique sont pour ainsi dire indispensables.

  • **Facilité de mise en œuvre et adaptabilité** : Un IDS doit être facile à mettre en œuvre et doit pouvoir surtout s'adapter au contexte (matériels, etc. . .) dans lequel il doit opérer.

- **Performances** : la mise en place d'un IDS ne doit en aucun cas affecter les performances des systèmes surveillés. De plus, il faut toujours avoir la certitude que l'IDS a la capacité de traiter toute l'information à sa disposition.
- **Multi-canal** : Un bon IDS doit pouvoir utiliser plusieurs canaux d'alerte (email, pager, téléphone, fax...) afin de pouvoir garantir que les alertes seront effectivement émises.
- **Information** : L'IDS doit donner un maximum d'information sur l'attaque détectée afin de préparer la réaction.
- **Classification** : il doit être aisé de hiérarchiser la gravité des attaques détectées afin d'adapter le mode d'alerte.

- **Quelques outils :** Voici quelques outils qui sont disponibles, on les distingue selon leur méthode de détection ainsi que sur leur modèle économique.

Nom de l'IDS	HIDS	NIDS	Comportementale	Scénario	Payant	Libre
Attack Mitigator		X		X	X	
Blackice		X		X	X	
Bro		X		X		X
Cisco IPS				X	X	
Dragon		X		X	X	
Prelude-IDS	X	X		X		X
SNORT		X	X	X	X	X

Choix

- SNORT est l'un des IDS les plus répondus et utilisés pour les raisons suivantes :

  – Actuellement, c'est le logiciel le plus répandu avec plus de 2000000 téléchargements.

  – Il bénéficie d'une mise à jour en temps réel (une version payante : OINKMAster via SourceFile et une autre gratuite : Bleeding via CERT), d'où la grande réactivité.

  – Il propose une architecture modulaire composée de :

    – Préprocesseurs.

    – Moteur de détection.

    – Système d'alerte et d'enregistrement de log.

    – Modules de sortie (pour enregistrer les logs dans des bases de données par exemple).

    – Décodeur de paquets