Détection des Intrusions, du monitoring des Systèmes d'Information au Graph Mining

David Pierrot, Nouria Harbi, Jérôme Darmont

▶ **To cite this version:**

David Pierrot, Nouria Harbi, Jérôme Darmont. Détection des Intrusions, du monitoring des Systèmes d'Information au Graph Mining. 4e Atelier International sur l'Innovation et Nouvelle Tendances dans les Systèmes d'Information (INTIS 2014), 2014, Rabat, Maroc. hal-01448382

HAL Id: hal-01448382
https://hal.science/hal-01448382v1

Submitted on 31 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Détection des Intrusions, du monitoring des Systèmes d'Information au Graph Mining

David PIERROT*, Nouria HARBI**, Jérôme DARMONT**

*Université Lumière Lyon 2, Laboratoire ERIC 69635 Lyon, Cedex FRANCE
david.pierrot1@univ-lyon2.fr
**{nouria.harbi,jerome.darmont}@univ-lyon2.fr

Résumé. La démocratisation d'Internet, couplée à l'effet de la mondialisation, a pour résultat d'interconnecter les personnes, les états et les entreprises. Le côté déplaisant de cette interconnexion mondiale des systèmes d'information réside dans un phénomène appelé "Cybercriminalité". Ainsi, des personnes, des groupes mal intentionnés ont pour objectif de nuire à l'intégrité des systèmes d'Information dans un but financier ou pour servir une "cause". Cependant, des moyens de protection existent depuis plusieurs années, mais ces derniers ne permettent pas une détection en temps réel et sont réservés aux seuls initiés. En conséquence, nous proposons une méthode d'analyse des flux en temps réel permettant de détecter les comportements anormaux et dangereux menaçant la sécurité des Systèmes d'Information et d'appréhender les risques d'une façon compréhensible par tous les acteurs.

1 Introduction

L'explosion d'internet, couplée à l'effet de la mondialisation, a pour résultat d'interconnecter les personnes, les entreprises, les états. Il devient donc plus aisé de communiquer, d'échanger, d'acquérir des biens, des connaissances via le réseau Internet. Selon le rapport du CLUSIF (2012) les entreprises ont une dépendance estimée à 81% par rapport à leur Système d'Information. Le maintien opérationnel d'un Système d'Information est devenu un des critères essentiels pour toute entreprise, administration ou personne cherchant à délivrer, proposer un service, ou simplement souhaitant communiquer. Le côté déplaisant de cette interconnexion mondiale des Systèmes d'Information réside dans un phénomène appelé "Cybercriminalité". Des personnes, des groupes mal intentionnés ont pour objectif de nuire dans un but pécuniaire ou pour une "cause", aux informations d'une entreprise, d'une personne voire d'un Etat. Il n'est pas rare que des faits de "cyber-attaques" soient relatés dans les médias envers des grandes sociétés comme "Google","Visa","Sony". La sécurité d'un Système d'Information se doit d'être présente afin de garantir la confidentialité, l'intégrité, la disponibilité de l'information. Conséquemment, la détection des intrusions et des activités liées à la "cybercriminalité" plus communément appelée Lutte Informatique Défensive (LID : M.Défense, 2008), doit permettre de protéger le Système d'Information. L'apparition de systèmes d'anonymisation comme "Tor" démontré par l'étude de Ujjaneni et Achutha (2013) et des outils comme "Shodan" (Sélégny

(2013)) ou "Google Hacking"(Lancor et Workman, 2007) limitant les traces laissées lors de la phase de reconnaissance (Akanksha, 2012) sur les éléments d'un Système d'Information, donne un sentiment d'impunité auprès de ses utilisateurs. Le métronome du rythme de la sécurité est bien donné par la communauté des "pirates informatiques" et des "Cybercriminels".
L'objectif de cet article est de présenter dans un premier temps l'état de l'art en matière de détection d'intrusion et dans un second temps d'aborder les travaux menés afin de faciliter la visualisation des flux. L'objectif final étant de fournir une appréciation du risque dynamique pour limiter l'exploitation des vulnérabilités.
La première partie de cet article sera consacrée à l'étude de l'existant dans laquelle nous présenterons les différentes approches de détection d'intrusion et leurs limites. Ensuite, on s'intéressera à la motivation de nos travaux et nous proposerons une solution. Nous détaillerons par la suite, la première phase de nos travaux ainsi que les résultats et nous terminerons par une conclusion et les perspectives.

2 Etude de l'existant

Les différentes avancées technologiques des dernières décennies ont pour effet une mise à jour pratiquement obligatoire de tout Système d'Information. Ainsi les "machines", les systèmes d'exploitation, les logiciels, les progiciels n'ont cessé d'être renouvelés. Il en va de même pour les "Hommes" qui ne peuvent échapper à un constant renouvellement de leurs connaissances et de leurs compétences pour maîtriser les nouvelles fonctionnalités des Systèmes d'Information. A partir de ce constat, tout Système d'Information est vulnérable par les éléments le constituant (logiciels, protocoles de communication, équipements d'infrastructures). Ainsi, une multitude d'outils (Antivirus, IDS, IPS, HIDS, Firewall) permettent de mettre en place une sécurité "relative" pour l'ensemble du Système d'Information. Un des principaux risques résiduels est d'obtenir un constat d'exploitation d'une vulnérabilité par une menace en temps réel. Il convient de répondre en fournissant des contremesures dans des délais raisonnables.

2.1 Les différentes solutions de détection d'intrusions

Les systèmes de détection des intrusions sont divisés selon les 3 familles distinctes :

— NIDS (Network Intrusion Detection System) est une sonde chargée d'analyser l'activité réseau du segment où elle est placée et de signaler les transactions anormales (Bhruyan et al (2011)).
— HDIS (Host-Based Intrusion Detection System) est basée sur l'analyse d'un hôte selon les produits utilisés, une HDIS surveille le trafic à destination de l'interface réseau, l'activité système et logiciel, les périphériques amovibles pouvant être connectés.
— HYBRIDES, ce type d'IDS rassemble les informations des NIDS et HIPS et produit des alertes aussi bien sur des aspects réseau qu'applicatifs.

Les IDS ont pour mission de détecter les tentatives d'intrusion, il existe aussi une variante appelée "IPS" (Intrusion Prevention System) étant capable d'appliquer une politique de sécurité lors d'une intrusion. Un IPS se décline à l'identique qu'un IDS, c'est à dire en mode réseau (NIPS) ou sur hôte (HIPS) ou encore en mode hybride. Un autre concept nommé CIDN ("Collaborative Intrusion Detection Networks") décrit par Fung (2011) offre la possibilité de

partager les informations relatives aux activités d'un Système d'Information entre différentes sondes déployées en son sein. Cette solution offre la possibilité de partager des informations entre des administrateurs d'un même réseau ou sur un espace communautaire sur Internet.
Les différentes solutions s'appuient sur deux méthodes qui sont utilisées par la plupart des "IDS" ("Open Source" et commerciales). La première méthode est fondée sur une comparaison d'une tentative d'intrusion par rapport à une base de signatures. Il s'agit de détecter les comportements abusifs. Ce type de système recherche dans les trames réseau un schéma qui correspond à une signature connue via l'utilisation de "Pattern Matching" (PCRE). Si la reconnaissance d'une activité réseau s'avère positive, un opérateur en charge de l'analyse et du traitement devra prendre en compte l'alerte. Il est aussi possible d'ajouter de nouvelles signatures, c'est à dire créer une expression régulière qui correspondra par son contenu à une activité malveillante ou abusive. Ce type d'IDS n'est pas dénué de sens pour les éditeurs et voire même pour la communauté "Open source ". Ainsi un modèle économique proche des éditeurs d'anti-virus peut faire bénéficier à une organisation, moyennant une redevance d'un système de détection à jour. Une des plus célèbres IDS nommée Snort est gratuite pour les signatures antérieures à un mois, mais une participation financière est demandée pour toute mise à jour inférieure à ce délai.
La seconde méthode est caractérisée par la détection des anomalies, elle repose sur des modèles comportementaux appelés "profils". Ils sont utilisés pour détecter tout comportement déviant des profils définis. Les anomalies peuvent " signaler" une intrusion ou un nouveau comportement. Dans le second cas, il convient d'ajouter ces nouveaux comportements afin de diminuer les " faux positifs". Le concept de détection des anomalies repose sur une analyse statistique et un apprentissage temporel des comportements. Plusieurs principes de mise en place sont disponibles comme "IDES : Intrusion-Detection Expert System" (Lunt et al, 1992), "NIDES : Next generation Intrusion Detection Expert System" (Anderson et al (1995), et Javitz et Valdes (1994)) ou "EMERALD : Event Monitoring Enabling Responses to Anomalous Live Disturbances" (Porras et Neumann, 1997).
Les deux systèmes évoqués présentent l'avantage d'être disponibles aussi bien dans le monde du logiciel libre que commercial, mais à une époque où la découverte et la gestion des vulnérabilités est devenue obligatoire pour la survie d'un Système d'Information, les inconvénients peuvent être considérés comme non négligeables. Le Tableau 1 liste les avantages et inconvénients des différentes solutions de détection.

2.2 Le Data Mining pour la détection d'intrusions

Depuis les premiers travaux du Docteur Denning (1987), les systèmes de détection n'ont cessé d'évoluer. Ainsi comme le précise Deepa et Kavitha (2012), le data mining propose différentes solutions pour détecter et analyser les attaques informatiques. Il est possible de créer des règles d'association (Amanpreet et al, 2011) ou d'utiliser les réseaux Bayésien (Kruegel et al., 2003 et, Huijuan et al., 2008) voire des arbres de décisions afin de detecter une intrusion. L'étude des Docteurs Sumeet et Xian (2011) montre les avantages du Data Mining face à une augmentation des données dans les Cyber-infrastructures et nombre croissant de cyberattaques. Dans un autre registre, le groupe "Thales" via son Laboratoire d'Innovation (Lagadec, 2012), propose une solution via une démarche d'analyse typologique des vulnérabilités du Système d'Information nommée "l'Analyse Topologique de Vulnérabilité". "TVA" a pour résultat des "graphes d'attaques". Cette solution permet la simulation d'attaque ce qui en fait un élément

	Avantages	Inconvénients
NIDS	Alarme en cas d'anomalie Positionnement multiple Temps réel	Signatures à mettre à jour Absorption du trafic Inopérant pour les flux chiffrés Gestion des faux positifs Expertise souhaitée
HIDS	Protège les stations Temps réel	Inefficace contre les attaques sur plusieurs hôtes Différentes configurations selon les systèmes utilisés
Hybride	Diminution des faux positifs Temps réel Corrélation des événements	Sources plus nombreuses, gestion et interprétation des alarmes plus difficiles

TAB. 1 – *Avantages et inconvénients des IDS "classiques".*

"proactif". Les graphiques générés permettent la définition et la mise en place d'une stratégie pour la prévention des attaques et limitent le plus possible les risques résiduels. Elle permet ainsi un déploiement des Sondes de détection d'Intrusions plus efficientes afin de couvrir les infrastructures sensibles et prioritaires. L'ensemble des chemins et des attaques potentielles étant connu, l'élimination des faux positifs est donc plus simple. Les graphiques donnent une vue horizontale du Système d'Information, ce qui a pour résultat une meilleure lisibilité et une diminution de l'interprétation des menaces propres aux métiers et processus prioritaires.

2.3 Limitation des solutions existantes

Les principales limites des outils présentés dans les chapitres précédents résident dans le fait qu'ils ne prennent pas encore en compte l'évolution quasi permanente d'un Système d'Information. Par exemple, la sécurité d'un entrepôt de données peut être mise en cause par la non réévaluation du ou des serveurs hébergeant ce dernier. La migration vers une nouvelle version du système peut remettre en cause la confidentialité ou l'intégrité voir la disponibilité des informations. La structuration organisationnelle et l'analyse de risques s'avèrent donc indispensables.

3 Motivations et proposition

La sécurité ne devant pas être réservée aux seuls experts, il convient de "vulgariser" les événements afin de les rendre compréhensibles par tous et d'automatiser au maximum les actions en découlant. La mise en place d'une supervision du Système d'information doit permettre d'analyser les flux de données de type événement et attaque, d'être en mesure de réagir en temps réel, et de corriger si nécessaire. L'addition de ces éléments doit être en mesure de prévoir les risques encourus par les différents actifs connus et suivis. Selon la déclinaison des motivations, notre étude porte sur quatre phases qui se décomposent de la façon suivante :

— Phase 1 : "Monitoring et visualisation" des données réseau, représentation graphique des activités d'un réseau informatique via un modèle de données.
— Phase 2 : "Extraction des profils d'attaques"
  Soit avec le Graph Mining ou des méthodes non supervisées (règles d'association) pour détecter des événements selon :
  — l'actif ciblé, les protocoles et services utilisés et selon les risques encourus
— Phase 3 : "Scoring" des risques et phase d'évaluation
— Phase 4 : Détermination d'un plan d'actions

3.1 Réalisation de la première phase

Il convient maintenant de présenter en détail la phase 1. Le principe est de modéliser un système de monitorage et de visualisation des données réseau en temps réel. Il est opportun de capturer les flux réseau à partir d'un équipement de filtrage "Firewall" et d'exporter les traces de connexion vers un conteneur de données. Par la suite, il est possible d'obtenir une vision graphique des flux transitant sur le réseau avec une rétroactivité de 3 jours.

3.1.1 Description des architectures

Il a été possible de tester la phase 1 sur plusieurs architectures et équipements de filtrage de marque et de technologie différentes. Le but étant d'obtenir une vision similaire dans des systèmes d'informations différents.

- Architectures de filtrage de second niveau

La première architecture concerne une entreprise publique dans le domaine de la santé composée de 125 000 employés. Notre étude porte sur 3 réseaux interconnectés au sein d'un reseau étendu (WAN : Wide Area Network) distant géographiquement. Ces derniers sont tous pourvus d'équipements de filtrages, l'objectif est d'être en mesure d'analyser les événements liés aux règles de filtrage via une exportation vers un conteneur de données. La figure 1 détaille sommairement l'architecture étudiée. Afin de simplifier les références aux différents réseaux, le nommage suivant sera utilisé :

— Site de production : **SP1**
— Site de qualification : **SQ1**
— Site d'administration distante et bureautique : **SAB1**

Ces trois sites sont opérationnels, c'est à dire que les données traitées et analysées dans les chapitres suivants correspondent à des données de production. Pour des raisons de **confidentialité** les adresses IP ont été anonymisées. Le réseau SP1 est doté de son propre conteneur de données issue des événements envoyés en temps réel par le Pare feu. Les réseaux SAB1 et SQ1 mutualisent un même conteneur. Les données bruts envoyées par l'ensemble des équipements filtrant sont traitées selon une reconnaissance de schéma. La phase 1 se focalise uniquement sur l'analyse et la représentation graphique de ces dernières.

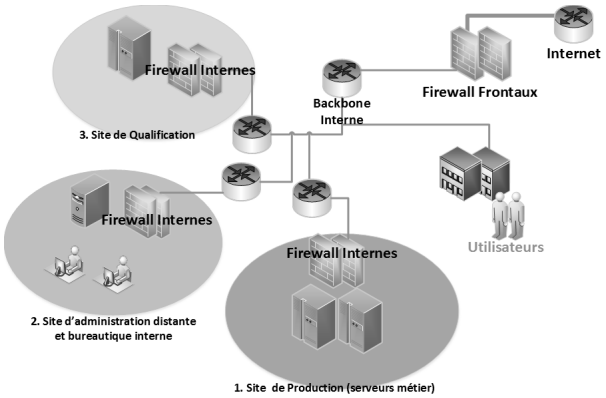

FIG. 1 – *Schéma du réseau étendu SP1,SQ1,SAB1*

- Architecture de filtrage frontal

Un second système d'information de taille plus modeste concerne une entreprise multinationale de 300 personnes dans le domaine du ludo-éducatif (tablette numérique, téléphone portable, jeux pour enfants). Cette dernière a pu être étudiée selon le même procédé. En revanche, il a été possible d'analyser les événements d'un "Pare feu" directement raccordé à Internet et servant de passerelle à l'ensemble des utilisateurs. L'équipement de filtrage étant de marque et de technologie différente, une modification du schéma de reconnaissance a du être réalisée. Ainsi un conteneur de données a pu recevoir l'ensemble des transactions réseau. Dans un but de simplification, le réseau étudié se nomme selon le terme suivant :

— Site frontal : **SF1**

La figure 2 représente de façon simplifiée l'architecture étudiée.

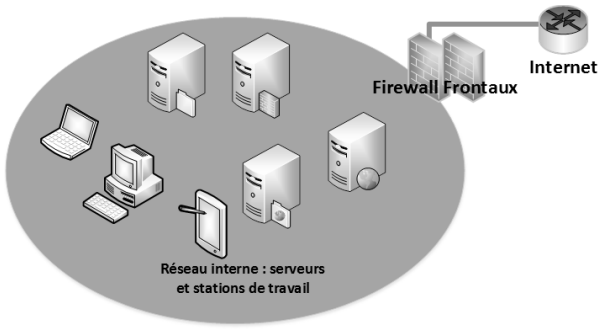

FIG. 2 – *Schéma d'architecture simplifié SF1*

3.1.2 Description des données

Le réseau SP1 propose des services à destination de **14 millions** de personnes. Les données peuvent être considérées comme sensibles et portent sur une quantité de 9.2 Teraoctets et plusieurs dizaines de millions d'euros par jours. Ces données sont hétérogènes et proviennent de plusieurs sources différentes. Les transactions réseau filtrées par le "Firewall" représentent plus de 6000 lignes par minute en matière d'événements. Le second et le troisième sites sont respectivement utilisés pour des taches dites de "bureautique-administration distante" et de qualification (reproduction des infrastructures de production).
Les modalités des variables listées ci-dessous sont exportées vers les conteneurs de données. La phase 1 se focalisera uniquement sur l'analyse et la représentation graphique de ces dernières.

- adresse ip source
- adresse ip de destination
- port de destination
- protocole (udp et tcp)
- date et heure de la connexion
- numéro de la règle du pare feu correspondant aux flux

Les tableaux 2 et 3 synthétisent le volume en nombre de lignes traitées par les équipements de filtrage.

	flux traités par journée	moyenne par minute
SP1	9 886 928	6 865
SAB1	572 272	397
SQ1	20 670	14

TAB. 2 – *Flux traités par SP1, SBQ1, SAB1*

	flux traités par journée	moyenne par seconde
SF1	716 045	497

TAB. 3 – *Flux traités par SF1*

Il s'agit par l'étude des métiers et surtout d'infrastructures différentes d'être en mesure de démontrer par le "monitoring graphique" que la compréhension des événements est plus facilement abordable et compréhensible. Pour résumer, le monitoring doit être un tableau de bord pour ses administrateurs. Il doit offrir la possibilité de répondre aux questions suivantes :

- "Que s'est-il passé sur mon réseau depuis hier soir ?".
- "Y a-t-il une augmentation de flux rejetés ?"
- "Le réseau est il victime d'une prise de renseignement ?"
- "Les règles de filtrage mises en place, sont elles suffisantes ?"

4 Exemple de réalisation

Le représentation graphique de l'ensemble des flux autorisés selon la période souhaitée peu paraitre "brouillone" (voir figure 3), mais il est possible de cibler uniquement les adresses source et de destination ainsi que les ports et protocoles souhaités, la lecture en sera que plus simple. En revanche, l'analyse d'un graphique fondé sur les flux rejetés (même agréés) comme le montre la figure 4 s'avère simple mais aussi efficace. Une adresse IP tente de se connecter à plusieurs autres adresses sur le port "135". Une recherche de répertoires partagés peut être à l'origine de ce type de comportement.

La figure 5 montre une seconde possibilités de représentation graphique basée sur le nombre

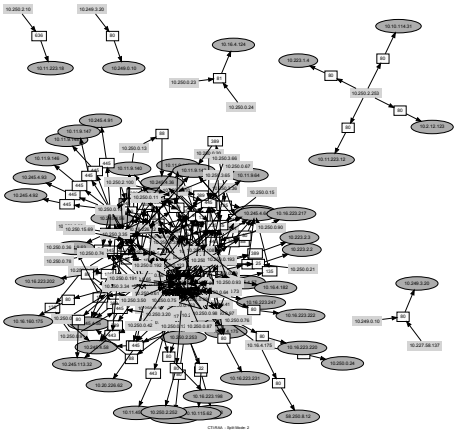

FIG. 3 – *Exemple de transactions*

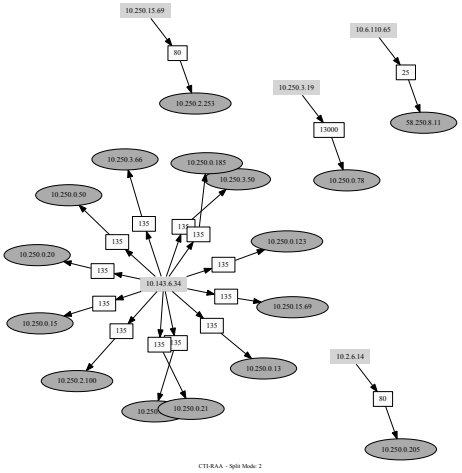

FIG. 4 – *Exemple de transactions rejetées*

de flux rejetés par tranches de 30 secondes. La variation étant relativement faible, les transactions réseau sont considérées comme normales.

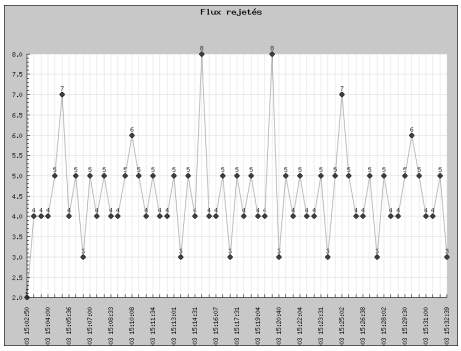

FIG. 5 – *Variation du nombre de flux rejetés par tranche de 30 secondes*

Par contre, si une activité similaire à la représentation graphique de la figure 6 est constatée, un balayage de ports est certainement à l'origine de cette affichage. La figure 7 peut être utilisée pour corroborer ce constat. Les modalités généralement constatées ont été largement outrepassées, passant de **9** rejets par tranche de 30 secondes à **3287**. Une surveillance peut être mise en place ainsi qu'une étude sur les menaces pouvant exploiter une vulnérabilité sur l'actif visé. Cette action est réalisable à condition d'avoir entrepris un inventaire des services et des vulnérabilités propres aux équipements (stations, serveurs, routeurs, etc...) (phases 2 à 4)

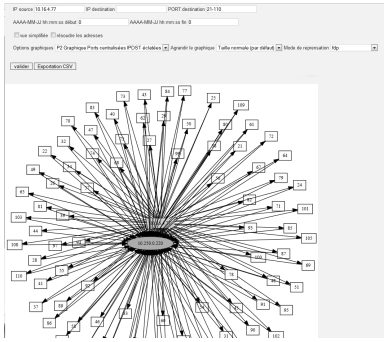

FIG. 6 – *Représentation graphique d'un balayage de ports (prise de renseignement)*

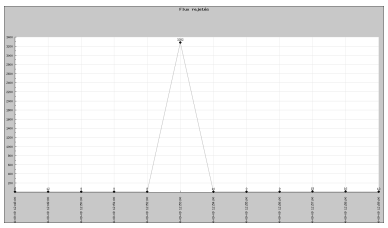

FIG. 7 – *Nombre de flux rejetés par tranche de 30 secondes lors d'une prise de renseignement*

4.1 Traitement et résultat

4.1.1 Traitement des données

Avant de pouvoir produire un résultat graphique et compréhensible, il convient de détailler les différentes opérations et configurations mises en place. Les différents moyens décrits dans ce chapitre permettent de réaliser le pré-traitement. Ces derniers dispensent à partir des événements bruts envoyés par les équipements de filtrage un format exploitable via le conteneur de données. La figure 8 présente en détail les différents étapes de conversions des données.
Les différentes traces de connexions des équipements de filtrage (option LOG) sont envoyées au serveur Syslog-NG. Ce dernier via ses options de reconnaissance PCRE et de filtrage dépose l'ensemble des flux traités sur un serveur de bases de données (MYSQL). Par la suite,

un traitement est réalisé via un programme Perl ayant pour objectif de préparer le résultat des différentes requêtes. Enfin, un second programme issue de la suite Graphviz est utilisé pour la création des graphiques. L'utilisateur n'a en faite besoin que d'un navigateur Internet pour être en mesure de visualiser les flux.

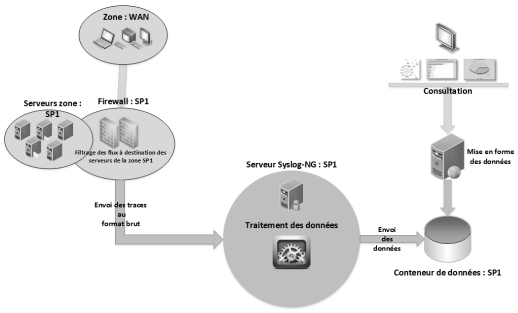

FIG. 8 – *Schéma traitement des événements*

4.1.2 Résultat

A l'issue de la phase 1, l'ensemble des événements lié au filtrage sont exportés en temps réel vers des conteneurs de données. Dans un souci de performance, seuls les trois derniers jours d'activités sont conservés. Le traitement des informations recueillies sur les différents équipements de filtrage permet une représentation graphique des flux. Par la suite, un critère de recherche fondé sur les flux rejetés a été mis en place. La représentation graphique du résultat obtenu permet de visualiser les activités anormales. La lecture graphique permet de visualiser rapidement les tentatives de connexion depuis plusieurs sources vers plusieurs destinations. Il peut s'agir d'une tentative d'intrusion ou d'une prise de renseignement ou encore d'une mauvaise configuration d'un script. Ce constat permet de soulever des interrogations sur cette transaction et de mettre en place une action de surveillance. Des alertes sont mises en place afin d'informer les administrateurs qu'une quantité de flux anormale a été rejetée (méthode manuelle qui sera complétée par les phases 2 à 4).
D'autres options ont été crées afin de permettre la visualisation des règles de filtrage les plus utilisées. En cas de doute sur une adresse Ip, il est possible de lister toutes les activités de cette dernière selon des critères de temps, de destination, de protocoles et de ports utilisés.
Afin d'assurer une réversibilité des traces, un module d'exportation a été ajoutée, ceci permet de faciliter les échanges entre différents intervenants en charge de la supervision et de la sécurité du Système d'Information. L'addition de ces différents modules permet une meilleur prise en charge de l'activité réseau pouvant laisser entrevoir un début de maitrise des risques induits et résiduels existants. Les prochaines phases de l'étude permettront de compléter ce point.

4.1.3 Résultat annexe

Le second système d'information **SF1** possédait à l'origine une politique de filtrage relativement faible. Seules 10 règles de filtrage étaient mises en place. De plus, certaines règles

s'avéraient trop permissives. Par l'analyse des graphiques de transaction et par le calcul des sommes des flux (non agrégées), il a été possible d'ajouter 80 nouvelles règles de filtrage dans un laps de temps très court et ceci avec un minium d'impact sur l'activité de l'entreprise. La phase 1 a permis de détecter des prises de renseignement, des tentatives de connexion à des ports destinés à l'administration distante à partir d'Internet. Ces phénomènes caractérisent généralement la préparation d'une intrusion.

5 Conclusions et perspectives

La réalisation de la première phase donne la possibilité de visualiser l'activité réseau et surtout les flux étant rejetés. Ces derniers donnent une vision de l'activité non autorisée et leur représentation graphique permet une compréhension aisée par rapport aux journaux d'événements bruts. L'étude s'est focalisée uniquement sur des informations de type session (adresse ip source, adresse ip des destinations, protocoles et services). Même si le résultat permet de faciliter l'ajout de règles de filtrage et un suivi des flux, le résultat peut être considéré comme **insuffisant** face aux nouvelles attaques et à l'augmentation constante du trafic réseau. Il convient de poursuivre les phases 2 à 4 à savoir l'extraction des profils d'attaques, établir un "scoring" et la création d'un plan d'actions basées sur l'application des méthodes de Data Mining sur un espace de représentation graphique. Ces phases permettront de générer des règles d'association en fonction des noeuds où une menace pouvant exploiter une vulnérabilité selon plusieurs vecteurs d'attaques. Une prise en compte des modes supervisés et non-supervisés permettrait aussi une meilleure prise en compte des attaques avec une définition automatique des seuils d'alertes. Il conviendrait de créer un système évolutif et adaptatif en temps réel permettant en fonction des différents changements intervenant sur un Système d'information de mettre automatiquement à jour l'évaluation des risques.

Références

Akanksha, B. (2012). Ethical hacking and social security. *Journal of Radix International Educational and Research Consortium 1*.

Amanpreet, H. et al (2011). Survey on data mining techniques in intrusion detection. *International Journal of Scientific Engineering Research 2*.

Anderson, D. et al (1995). Next-generation intrusion-detection expert system (nides). Technical report, SRI International, Menlo Park, California.

Bhruyan, H. et al (2011). Survey on incremental approaches for network anomaly detection. *International Journal of Communication Networks and Information Security 3*.

CLUSIF (2012). Menaces informatiques et pratiques de sécurité en france. Technical report, CLUSIF (https ://www.clusif.asso.fr/).

Deepa, A. J. et V. Kavitha (2012). A comprehenseive survey on approaches to intrusion detection system. *ICMOC-2012*.

Denning, D. (1987). An intrusion-detection model. *IEEE Transactions on Software Engineering*.

Fung, C. (2011). Collaborative intrusion detection networks and insider attacks.

Graphviz. Graphviz - graph visualization software, http://www.graphviz.org.

Huijuan, L. et al. (2008). Two stratum bayesian network based anomaly detection model for intrusion detection system.

Javitz, H. S. et A. Valdes (1994). The nides statistical component description and justification. Technical report, SRI International, Menlo Park, California.

Kruegel, C. et al. (2003). Bayesian event classification for intrusion detection.

Lagadec, P. (2012). Visualisation et analyse de risque dynamique pour la cyber-défense.

Lancor, L. et R. Workman (2007). Using google hacking to enhance defense strategies. *ACM SIGCSE Bulletin 39*.

Lunt, T. F. et al (1992). A real-time intrusion-detection expert system (ides). Technical report, SRI International, Menlo Park, California.

M.Défense (2008). Bulletin officiel des armées edition chronolique n°44. Technical report, Ministère de la Défense (http ://www.boc.sga.defense.gouv.fr/).

MYSQL. The world's most popular open source database, http ://www.mysql.com.

PCRE. Perl compatible regular expression, http ://www.pcre.org.

Perl, t. T.

Porras, P. A. et P. G. Neumann (1997). EMERALD : event monitoring enabling responses to anomalous live disturbances. In *1997 National Information Systems Security Conference*.

Sélégny, G. (2013). Shodan, un moteur de recherche alimente des peurs justifiées. Technical report, Centre National de ressource et d'information sur l'Intelligence économique et stratégique (http ://www.portail-ie.fr/article/834/Shodan-un-moteur-de-recherche-alimente-des-peurs-justifiees).

Snort. Open source network intrusion prevention and detection system, http ://www.snorg.org.

Sumeet, D. et D. Xian (2011). *Data Mining and Machine Learning in Cybersecurity*.

Syslog-NG. Log management, http://www.balabit.com/network-security/syslog-ng.

Ujjaneni, S. et S. V. Achutha (2013). A novel cell reckoning intrusion against tor. *International of Computer Trends and Technology IJCTT 4*.

Summary

The democratization of the Internet, coupled with the effect of globalization, resulting interconnect personns, states and companies. The unpleasant side of this global interconnection of information systems is called "Cybercrime". Thus, individuals, groups malicious aim to undermine the integrity of information systems for financial gain or to serve as a "cause". However, means of protection exist for several years, but they do not allow real-time detection and reserved for the initiated. Accordingly, we propose a method for analyzing real-time streams to detect abnormal behavior and dangerous threatening the security of information systems and understand the risks in an understandable by all stakeholders.