*Afrique* SCIENCE *13(4) (2017) 310 - 326*
ISSN 1813-548X, http://www.afriquescience.info


# Sécurisation réseau à partir de la théorie de graphe

**Rado RAZAFY*, Soloniaina RAKOTOMIRAHO et Rivo RANDRIAMAROSON**

*Université d'Antananarivo, Ecole Supérieure Polytechnique, BP 1500, Antananarivo 101, Madagascar*

* Correspondance, courriel : *r.razafy@bfm.mg*

## Résumé

L'évolution technologique force le monde à partager des informations ou à accéder à des informations partagées. « S'isoler » n'est pas une solution pour « se protéger » contre les menaces. La mise en place d'une sécurité informatique doit commencer par l'analyse des risques. La théorie de graphe a été prise dans cet article pour la réalisation de l'analyse. Les trois types d'attaque, qui sont les attaques d'accès et vols d'information, les attaques de saturation et les attaques de répudiation, sont étudiés à partir de cette théorie. Le résultat de l'étude a permis de dégager une architecture respectant les mesures de sécurité. Principalement les solutions de regroupement et la gestion des flux sont identifiées. La théorie de graphe est aussi identifiée comme un outil pour l'implémentation et le déploiement des équipements. Une amélioration de l'architecture est nécessaire pour minimiser les risques.

**Mots-clés :** *sécurité informatique, modélisation réseau, théorie de graphe, analyse de risque.*

## Abstract

### Network security based on graph theory

The evolution of Information Technology is forcing the whole world to share information or access to shared information. Being "isolated" or "in our own" is no longer a solution that could be considered "to protect" oneself against any threats. An implementation of an efficient information security has to begin with a risk analysis. This article, especially the "analysis" part is entirely based on the Graph Theory. Three types of attacks are studied throughout this article : access attack and information theft, stressing attack and repudiation attack. The outcomes of the study allowed us to define an architecture that respects all security measures (described herein this document). Data flow management and aggregation solutions are identified. The Graph theory comes out also as an efficient tool for the network nodes implementation and deployment. However, the overall architecture requires an improvement to minimize the risks.

**Keywords :** *information security, network architecture, graph theory, risk analysis.*

## 1. Introduction

En mai 2017, le virus « Wanna Cry » ou « Wanna Crypt » de type « rancomware » a touché dans 150 pays et 200 000 victimes [1]. Pour quelque référence, Renault en France, FedEx en Etats Unis ont subi cette attaque et

ont eu d'impact sur leur exploitation [2]. Toutes leurs informations et toutes leurs productions sont actuellement informatisées. En 2016, un hôpital a été aussi paralysé par un tel type d'attaque [3]. Le virus a chiffré les données sur le disque dur et rend illisibles à leur propriétaire. Tous les fichiers médicaux des patients et les fichiers d'admission de l'hôpital sont infectés par le virus. De plus le pirate a pu perturbe le réseau pour rendre indisponibles les équipements électroniques. Cette illustration montre qu'aucune institution n'est à l'abri d'une attaque. En avril 2016, certaines banques américaines ont été une cible de DDoS qui ont à ternir leur image [4]. D'après l'analyse de la société Prolexic, spécialisée dans la prévention des attaques par déni de service, le hacker a utilisé un outil qui est baptisé « itsoknoproblembro ». Qui pense qu'un fichier PDF présente un danger ? Un pirate peut facilement utiliser Metasploit pour pouvoir créer un PDF malveillant [5]. L'utilitaire permet d'insérer des codes qui vont être exécutés à l'ouverture du fichier. A partir de ces codes le pirate peut accéder à la machine pour soustraire des informations ou pour explorer le réseau contenant l'ordinateur. A partir d'une telle attaque, le pirate peut prendre contrôle l'installation entière. En ayant accès sur le réseau, il peut commencer à voir et à contourner les autres failles. D'après l'analyse « The Hacker News », cette pratique a été utilisée pour le piratage de la Banque Centrale de Bangladesh [6]. Finalement le pirate appris contrôle l'installation de la banque. Les menaces informatiques sont les actions ou les évènements, volontaires ou non, pouvant nuire au bon fonctionnement d'un système d'information. La sécurité des systèmes d'information est l'ensemble des moyens techniques, organisationnels, juridiques et humains nécessaires mis en place pour conserver, rétablir, et garantir la sécurité du système d'information. La modélisation de réseau assurant la sécurité a pour but de mettre en place ses mesures au niveau de la liaison qui assure l'échange des informations utiles mais aussi qui joue le rôle de la propagation des attaques. La théorie de graphe a été prise comme théorie mathématique de base pour la modélisation. L'objectif est d'identifier les vulnérabilités et aussi de minimiser les risques.

## 2. Méthodologie

### 2-1. La théorie des graphes

La théorie des graphes fait partie du domaine de la mathématique découvert en 1736. Cette science naquit quand Euler avait démontré qu'il était impossible de traverser chacun des sept ponts de la ville russe de Königsberg (Kaliningrad) une fois exactement et de revenir au point de départ ***(Figure 1)***. Les ponts enjambent les bras de la Pregel qui coulent de part et d'autre de l'île de Kneiphof [7].

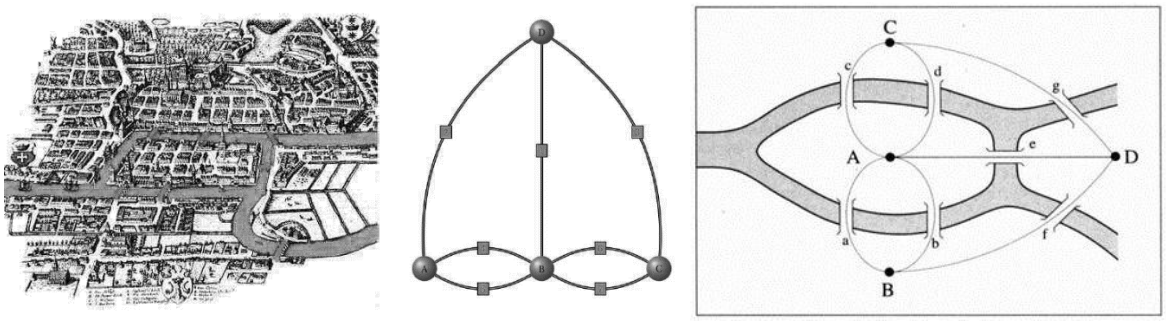



**Figure 1 :** *Pont de Königsberg*

Un graphe fini G = (V, E) est défini par l'ensemble fini V= $\{v_1, v_2, \ldots, v_n\}$ dont les éléments sont appelés « sommet » ou « nœud », et par l'ensemble fini E = $\{e_1, e_2, \ldots, e_m\}$ dont les éléments sont appelés « arêtes ». Une arête e de l'ensemble E est définie par une paire non ordonnée de sommets, appelés les extrémités de e. Si l'arête e relie les sommets a et b, on dira que ces sommets sont adjacents ou incidents avec e, ou bien

que l'arête e est incidente avec les sommets a et b. On appelle « ordre » d'un graphe le nombre de sommets n de ce graphe [8]. En donnant un sens aux arêtes d'un graphe, le graphe devient un digraphe (tiré de l'expression en anglais directed graph) ou graphe orienté. Un ***(Figure 2)*** digraphe fini $G = (V, E)$ est défini par l'ensemble fini $V = \{v_1, v_2, \ldots, v_n\}$ dont les éléments sont appelés « sommets », et par l'ensemble fini $E = \{e_1, e_2, \ldots, e_m\}$ dont les éléments sont appelés « arcs »[8].

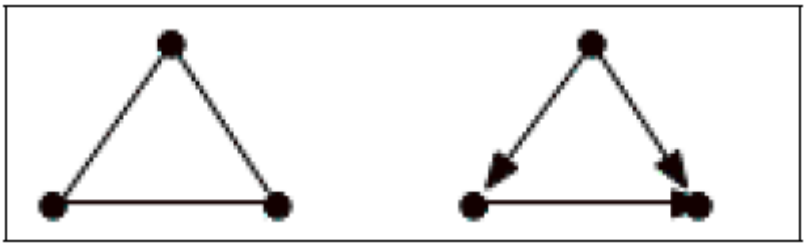


**Figure 2 :** *Graphe simple et graphe orienté*

L'adjacence d'un graphe peut être représentée avec une matrice. Soit $A(G) = a_{ij}$, la matrice d'adjacent du graphe G, dont $a_{ij}$ est le nombre d'arêtes ou d'arcs joignant les sommets $v_i$ et $v_j$. Elle fait la correspondance entre les sommets (ligne) origine des arêtes aux sommets destination (colonne). Dans le formalisme matrice booléenne, l'existence d'un arc $(x_i, x_j)$ se traduit par la présence d'une intersection de la ligne $x_i$ et de la colonne $x_j$. Dans le cas d'absence d'arc, on met la valeur 0. La ***Figure 3*** représente un exemple de la représentation. Dans le cas d'une matrice non orienté, les valeurs des arêtes peuvent être placées dans les intersections ***(Figure 4)***.

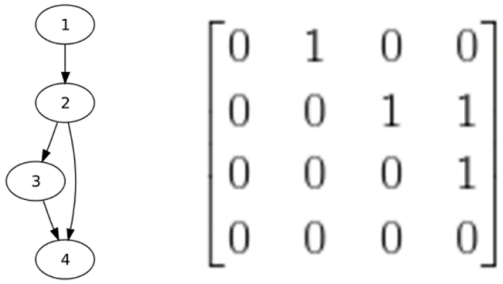



**Figure 3 :** *Exemple de représentation d'un graphe orienté*

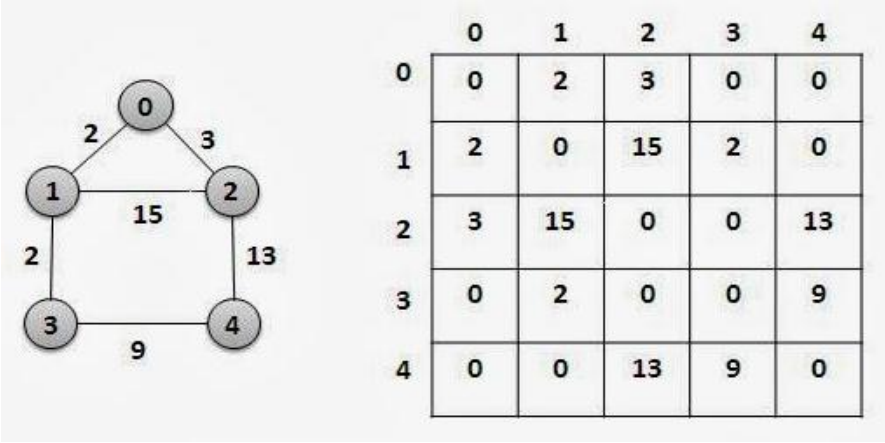



**Figure 4 :** *Exemple de représentation d'un graphe simple*

Dans la théorie des graphes, une chaine est une séquence d'arcs telle que chaque arc ait une extrémité commune avec le suivant. Un cycle est une chaine qui contient au moins une arête, telle que toutes les arrêtes de la séquence soient différentes et dont les extrémités coïncident. Un chemin et un circuit ont les mêmes définitions que la chaine et le cycle mais dans le concept de graphe orienté. Le sous-ensemble de sommets atteignables à partir d'un sommet donné, grâce à des chemins, est appelé fermeture transitive de ce sommet. Le terme parcours regroupe les chemins, les chaines, les circuits et les cycles.

### 2-2. Les principales définitions

- Adjacence :

Deux sommets sont adjacents (ou voisins) s'ils sont joints par un arc. Deux arcs sont adjacents s'ils ont au moins une extrémité commune.

- Degrés :
  - Le demi-degré extérieur de $x_i$, $d^+(x_i)$, est le nombre d'arcs ayant $x_i$ comme extrémité initiale ; $d^+(x_i) = |\omega^+(x_i)|$ ;
  - Le demi-degré intérieur de $x_i$, $d^-(x_i)$, est le nombre d'arcs ayant $x_i$ comme extrémité finale ; $d^-(x_i) = |\omega^-(x_i)|$ ;
  - Le degré de $x_i$ est $d(x_i) = d^+(x_i) + d^-(x_i)$. Le degré d'un sommet d'un graphe non orienté est le nombre d'arêtes qui lui sont incidentes.

- Graphe complémentaires :

$G = (X, U)$ et $\overline{G} = (X, \overline{U})$. $(x_i, x_j) \in U \Rightarrow (x_i, x_j) \notin \overline{U}$ et $(x_i, x_j) \notin U \Rightarrow (x_i, x_j) \in \overline{U}$. $\overline{G}$ et le graphe complémentaire de G.

- Graphe partiel :

$G = (X, U)$ et $U_p \subset U$. $G_p = (X, U_p)$ est un graphe partiel de G.

- Sous- graphe :

$G = (X, U)$ et $X_s \subset X$. $G_s = (X_s, V)$ est un sous-graphe de G, où V est la restriction de la fonction caractéristique de U à $X_s$. $V = \{(x, y) / (x, y) \in U \cap X_s \times X_s\}$. $\forall x_i \in X_s, \Gamma_s(x_i) = \Gamma(x_i) \Lambda X_s$.

### 2-3. Vocabulaire lié à la connexité

- Voisin ou contact :

Soit $u$ un sommet d'un graphe orienté $G = (V, E)$, tout sommet $v \in V$ tel que $(u, v) \in E$ est appelé voisin ou contact de $u$.

- Chemin :

Soit $G = (V, E)$ un graphe orienté et deux sommets $u, v \in V$, le *chemin* de $u$ vers $v$ de *longueurl* est toute suite $x_0, x_1, \ldots, x_{l-1}, x_l$ où $x_0 = u, x_l = v$ et $\forall\, i = [0, l-1], (x_i, x_{i+1}) \in E$.

- Distance :

Soit $G = (V, E)$ un graphe orienté et deux sommets $u, v \in V$, la *distance* de $u$ vers $v$ est la *longueur* du plus court *chemin* de $u$ vers $v$.

- Diamètre :

Le *diamètre* d'un graphe G est la plus grande des *distances* entre les pairs sommets de G.

### 2-4. Routage

- Graphe augmentée :

Un graphe augmenté $G = (V, E, E')$ est un graphe obtenu à partir d'un graphe $H = (V, E)$, en ajoutant un ensemble d'arrêt supplémentaires $E'$ sur $V$. La distance sous-jacente de $u à v$ dans $G$ est la distance de $u à v$ dans $H$.

- Algorithme de routage :

Soit $G = (V, E)$ un graphe orienté et deux sommets $s, t \in V$, l'algorithme de routage A dans le graphe G ayant les sommets $sourcesetcible(oudestination)t$ en entrée et renvoie en sortie un chemin P de $s$ à $t$. P est appelé chemin de routage de $s$ à $t$. L'efficacité d'un algorithme de routage décentralisé se mesure par la longueur des chemins calculés et la latence qui est le nombre de nœuds visités ou interrogés. Durant le calcul de chemin, l'algorithme peut interroger des nœuds qui ont d'informations utiles sans que ces nœuds appartiennent au chemin renvoyé par l'algorithme. Soit $H = (V, E)$ sous-jacent à un graphe augmenté $G = (V, E, E')$, un algorithme de routage dans un graphe augmenté $G$ est décentralisé si et seulement si il a pour seule connaissance le graphe sous-jacent $H$, la position de la cible sur $H$ et peut seulement interroger les nœuds connus ou les voisins de nœuds connus.

- Latence :

La latence d'un algorithme de routage sur un graphe donné est le plus grand nombre de nœuds interrogés pour calculer un chemin entre deux sommets.

## 3. Analyse des risques à partir de la théorie de graphe

### 3-1. Présentation d'un réseau

Un réseau non organisé est une image d'un graphe aléatoire. Le modèle d'Erdös et Rényl peut être pris en référence pour illustrer un tel type de graphe. Ce modèle consiste en un ensemble de n nœuds reliés par des arêtes qui sont placées de manière aléatoire uniforme entre paire de nœuds. Le plus communément étudié est celui dénommé $G_{n,p}$ où chaque arête entre deux nœuds est indépendamment présenté avec une probabilité $p$ et absente avec une probabilité $1 - p$. $G_{n,p}$ est caractérisé par $z$ qui est le degré moyen de z des nœuds. Le nombre moyen d'arêtes est $\frac{1}{2}\ n\ (n-1)p$ et le moyen de connexion correspond à 2 fois ce résultat puisque chaque arrêt a deux extrémités [9]. Le degré moyen des nœuds est donné par ***l'Équation 1.***

$$z = \frac{n\,(n-1)p}{n} = (-1)p\ \cong np \qquad (1)$$

Les équipements connectés peuvent être considérés comme des sommets et les liaisons comme des arêtes. Dans le cas de nécessité de considérer le sens des flux, on peut utiliser la notion d'arc. La ***Figure 5*** suivante montre la complexité de schématisation d'interaction pour un réseau social.

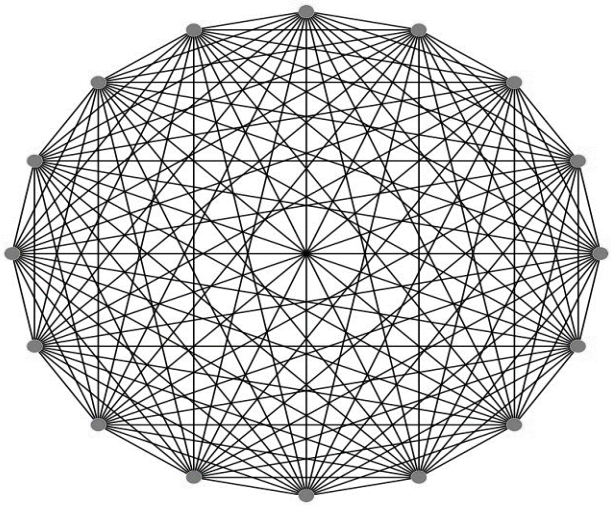


**Figure 5 :** *Exemple de schématisation de connexion pour un réseau social*

Même pour un simple système d'information au sein d'une entreprise, une modélisation reste complexe. Dans un réseau plat, tous les équipements peuvent se communiquer. Ce type de réseau est comme un graphe fortement connexe. Un graphe G = (X, U) est fortement connexe si $\forall\ i, j \in X$, il existe un chemin entre i et j. La vitesse de propagation de virus, les attaques de saturation et la reconnaissance effectuée par des pirates sont difficiles à surveiller dans un tel type de réseau. En effet, il existe plusieurs chemins pour arriver à un équipement. Tous les équipements peuvent s'échanger des informations entre eux et la mise en place des règles de sécurité devient impossible. En absence d'une étude de flux, un réseau ressemble à un graphe complet. Un graphe est complet, ***Figure 6***, quand $\forall x_i, x_j \in X, (x_i, x_j) \notin U \Rightarrow (x_i, x_j) \in U$. Dans un tel type de graphe, le degré moyen des nœuds est z = n.

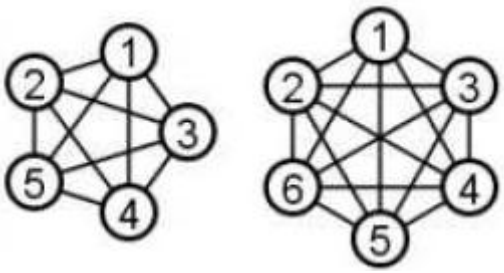



**Figure 6 :** *Exemple de graphe complet avec 5 et 6 sommets*

### 3-2. Les attaques d'accès et vols d'information

L'établissement des chemins dans un réseau est assuré essentiellement par le protocole ARP et la table CAM. Le protocole ARP (AdressRouting Protocole) et les tables CAM (Content Adressable Memory) n'interviennent pas sur les mêmes couches du modèle OSI. En effet, le protocole ARP assure la jointure entre la couche 3 (couche réseau, adresse IP) et la couche 2 (couche liaison, adresse MAC), alors que les tables CAM le font entre la couche 2 (liaison, adresse MAC) et la couche 1 (physique et numéro de port physique du switch) [10]. Pour router le trafic, les switch maintiennent la table CAM. Les chemins entre les nœuds sont donc assures par les switch (Si les switch fonctionnent correctement). Soit un graphe G = (X, U) une représentation du réseau, $\forall\ i, j \in X$, le chemin entre $i$ et $j$ est établie à partir de la table CAM. Soit $z \in X$, le chemin entre $i$ et $j$ ne passera pas par $z$. Le réseau va se simplifier ***(Figure 7)*** et ne ressemblera plus par un graphe complet.

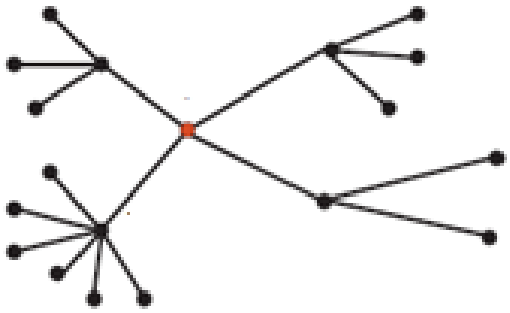


**Figure 7 :** *Représentation d'un réseau avec des switch*

Il est possible d'exploiter un dépassement de mémoire de certains équipements en saturant la table d'adresse MAC, afin de capturer les paquets. En effet, lorsque la mémoire est pleine, des switch se comportent en hub et envoient les paquets à tous les hôtes connectés au réseau [10]. Dans ce cas le réseau va se modéliser comme un graphe complet. Tous les équipements vont recevoir les informations circulant sur le réseau. En modifiant la configuration de la carte réseau en mode promiscuité, un pirate peut utiliser un logiciel d'écoute pour analyser et soustraire des informations confidentielles sur le réseau. A part l'utilisation de technique de sniffing, le risque d'exploration par un scanner de vulnérabilité est aussi élevé dans un tel type de réseau. Soit un graphe G = (X, U) fortement connexe représente un réseau plat. $\forall\ i, j \in X$, il existe un chemin entre i et j. un ordinateur ayant un utilitaire de vulnérabilité peut donc analyser tous les équipements du réseau. Le programme de sonde envoi sur un réseau de données et analyse les réponses reçus. Le programme fait un balayage de tous les ports afin de déterminer leur comportement. Un pirate peut perpètre des attaques en ayant la cartographie complète du réseau et les composants.

### 3-3. Les attaques par saturation

Les attaques par saturation peuvent se manifester avec trois formes qui sont :

- ✓ attaque d'une cible à partir de plusieurs sources ;
- ✓ attaque à partir d'une source vers un ensemble de cible ;
- ✓ attaque par rebond.

Les charges d'une cible quelconque dans l'ensemble du réseau peuvent être calculées à partir du degré moyen des nœuds. Le degré moyen des nœuds est donné par ***l'Équation 2.***

$$z = \frac{n\,(n-1)p}{n} = (-1)p \cong np \qquad (2)$$

Soit un graphe G = (X, U) une représentation d'un réseau plat, soit $i$ est une cible dans le réseau $\forall\ j \in X$, il existe un chemin entre $i$ et $j$. Le degré moyen des nœuds z = n avec n le nombre de nœuds. Pour une attaque de saturation, le pirate exploite tous les équipements du réseau pour envoyer des commandes vers la cible. Supposons que chaque équipement envoi r requêtes par seconde vers la cible, la charge supportée par la cible sera donc $nr$. Par exemple, le TCP-SYN flooding est une variante de flooding qui s'appuie sur une faille du protocole TCP. L'attaque se trouve au niveau 4 (transport) de la couche TCP / IP. Le principe est d'envoyer un grand nombre de demandes de connexions au serveur (SYN) à partir de plusieurs machines. Le serveur va retourner et répondre avec le paquet SYN-ACK et attendre en retour une réponse ACK qui n'arrivera jamais.

A la saturation de la queue permettant de stocker les connections en attente de fin d'ouverture, la machine n'acceptera plus aucune nouvelle demande. Par exemple, dans un réseau avec 100 postes clients et un serveur comme cible, si chaque poste de travail envoi simultanément $r$ demande de connexion par seconde, le serveur doit satisfaire $r * 100$ demandes de connexion par seconde. Dans un autre cas, une attaque peut venir d'une seule source vers l'ensemble de réseau. Soit un graphe G = (X, U) fortement connexe représente un réseau plat, soit $p \epsilon X$ la source de l'attaque, $\forall\ i \epsilon X$, il existe un chemin entre $p$ et $i$. La dernière alternative d'attaque possible consiste à faire des attaques par rebond qui est la combinaison des deux attaques précédentes. Le principe est de faire passer les requêtes frauduleuses à travers un équipement du réseau.

### 3-4. Les attaques de répudiation

La répudiation est une attaque contre la responsabilité. Autrement dit, elle consiste à tenter de donner de fausses informations ou de nier qu'un évènement ou une transaction se soit réellement passé. C'est le cas de l'IP spoofing appelé aussi mystification, il n'est pas une attaque en tant que tel mais une méthode pour dissiper l'identité. La méthode est de masquer l'origine d'une action en falsifiant l'adresse IP. La source se communique avec la machine cible comme une machine de confiance. L'attaque par rebond est une méthode pour cacher l'identité et faire passer l'attaque par une autre machine. L'objectif est de masquer les traces permettant de remonter jusqu'à la machine source. Dans le cas d'un réseau sous forme de graphe aléatoire, il est impossible de mettre des contrôles d'identité. En outre, comme tous les équipements peuvent se communiquer entre eux les pirates peuvent utiliser d'autre équipement en tant que pivot pour camoufler leur identité.

## 4. Amélioration de l'architecture

### 4-1. Regroupement par communauté

Pour réduire la complexité, les algorithmes de « communauté » comme « waltrap » et « fastgreed » peuvent être utilisés pour constituer des groupes distincts. Le premier est basé sur une probabilité proportionnelle aux poids des arêtes et aux degrés des sommets voisins. Le second se base sur une décomposition hiérarchique des arêtes du graphe pour mettre en évidence les communautés [11]. La ***Figure 8*** est un exemple de représentation des communautés dans un graphe.

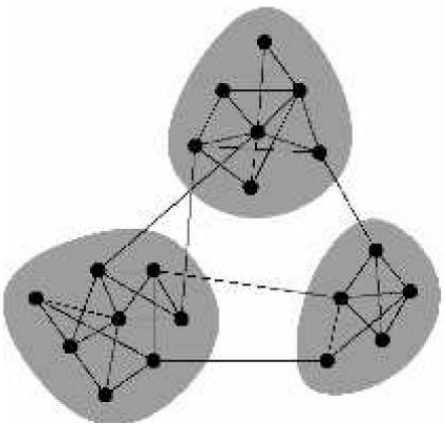
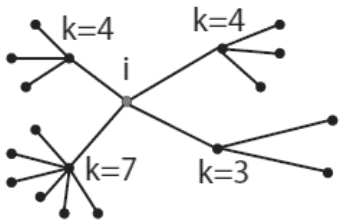


**Figure 8 :** *Communautés dans un graphe*

Le degré de corrélation ou le degré moyen des voisins d'un sommet i est obtenu par ***l'Équation 3***.

$$k_{nn,i} = \frac{1}{k_i}\sum_j a_{ij}k_j \quad (3)$$

Voici un exemple pour le graphe su la ***Figure 9***.




**Figue 9 :** *Exemple de graphe pour calculer le degré de corrélation*

$$k_i = 4$$
$$k_{nn,i} = \frac{1}{4}(3+4+4+7) = 4{,}5$$

Les sous-graphes dans les communautés peuvent être aussi obtenus par la « méthode de clustering » ou par la détection de sous-groupes par k-cores (Seidman, 1983). La première méthode consiste à la recherche les entités partageant les mêmes relations ou ayant les mêmes propriétés. Le second principe se base sur le degré des nœuds. Un k-core est un sous-graphe dont chaque somme est adjacent à un minimum de k autres sommets du sous-graphe. Pour appliquer ces principes dans une architecture réseau, les sous-graphes ou les communautés peuvent effectuer en faisant une séparation des réseaux. Dans le pratique, on utilise le technique de réseau virtuel ou VLAN. Par exemple, les réseaux les différentes directions ou aussi les départements dans une entreprise peuvent être classés par VLAN. Soit un graphe $G = (X, U)\ et\ X_s\ C\ X. G_s = (X_s, V)$ est un sous graphe de G, où V est la restriction de la fonction caractéristique de U à $X_S$. $V = \{(x,y)/(x,y) \in U \cap X_s \times X_S\}. \forall x_i \in X_s, \Gamma_s(x_{i)}) = \Gamma_s(x_{i)})\ \Lambda\ X_s.$

### 4-2. Réduction de la surface d'attaque par segmentation réseau

Vu précédemment, un réseau ayant comme représentation un graphe complet est vulnérable pour toutes les attaques. Le réseau dans une entreprise est généralement lié physiquement. Les postes de travail, les serveurs, les imprimantes ont les mêmes adresses ***(Figure 10)***. Un tel type de réseau est semblable à un graphe complet. En effet, les équipements peuvent s'échanger entre eux. La réduction de la surface d'attaque consiste à séparer les équipements pour éliminer les échanges inutiles et pour mieux identifier les sources d'attaque.

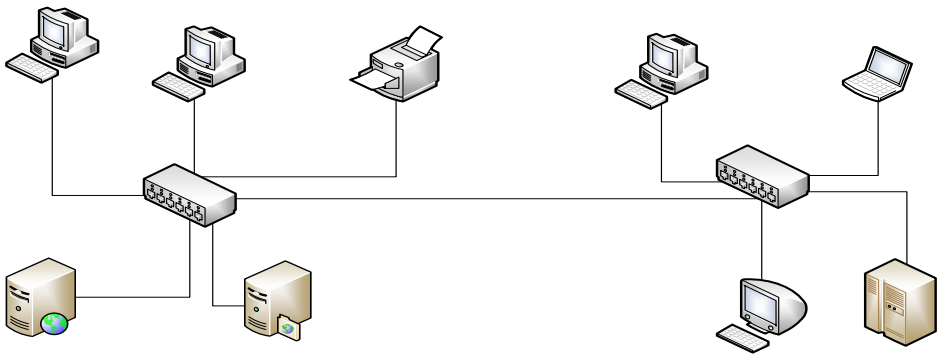


**Figure 10 :** *Exemple classique d'un réseau d'entreprise*

La répartition des équipements sur différents switch ne sépare pas logiquement les équipements. En effet, ayant les mêmes adresses, ils peuvent se communiquer entre eux sans aucune restriction. Tous les hôtes reçoivent les messages de diffusion. Un pirate peut faire une écoute du trafic échangé sur le réseau pour prendre les informations nécessaires en vue d'une préparation d'action malfaisante comme les attaques par saturation, déploiement d'un virus ou vols d'information. Un pirate peut exploiter la faille des switch qui maintiennent la table CAM pour que toutes les informations soient envoyées à tous les équipements. Il est possible d'exploiter un dépassement de mémoire de certains équipements en saturant la table d'adresse MAC, afin de capturer les paquets. Par définition, un réseau local est délimité par une interface d'équipement de niveau 3 du modèle OSI. Il est impossible de définir dans ce cas une règle de sécurité pour l'appliquer à une partie du réseau. La première méthode consiste à utiliser une séparation physique avec un routeur ***(Figure 11)***.

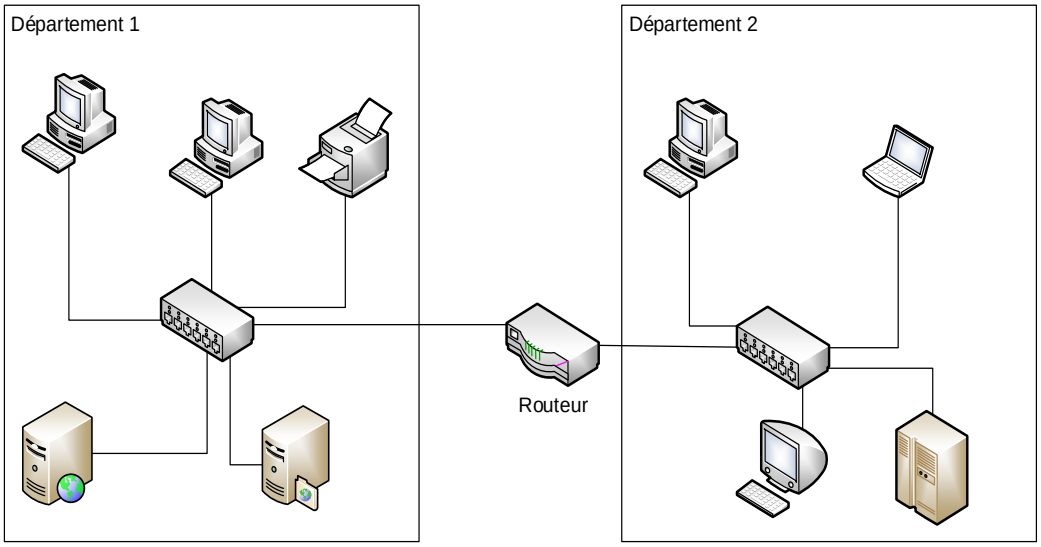



**Figure 11 :** *Représentation d'un réseau séparé physiquement*

Dans ce cas, les équipements de deux départements ont différentes adresses. Les sous réseaux peuvent avoir des règles différentes et pour aller d'un équipement du « Département 1 » vers le « Département 2 », il est obligatoire de passer par le routeur. Pour un système d'information, la mise en place d'une telle architecture est couteuse et parfois difficile car les utilisateurs sont difficiles à classer par leur emplacement physique. Le concept de réseau virtuel ou VLAN comble la limitation de la séparation physique. Il permet de faire une séparation même pour un équipement connecté sur un équipement seul ***(Figure 12)***. La configuration peut être aussi utilisée même pour les entreprises dont leurs installations se trouvent dans des localités différentes en utilisant le « Trunking » ***(Figure 13)***. L'objectif d'un VLAN est de réduire les surfaces d'attaque et d'offrir la possibilité d'appliquer les mesures de contrôle. Ainsi, la mise en place d'un VLAN exploite les protocoles utilisés sur les trois premières couches OSI. A noter qu'il existe trois types de VLAN qui sont VLAN de niveau 1, VLAN de niveau 2 et VLAN niveau 3. La norme 802.1q est la norme suivi pour l'établissement d'un VLAN.

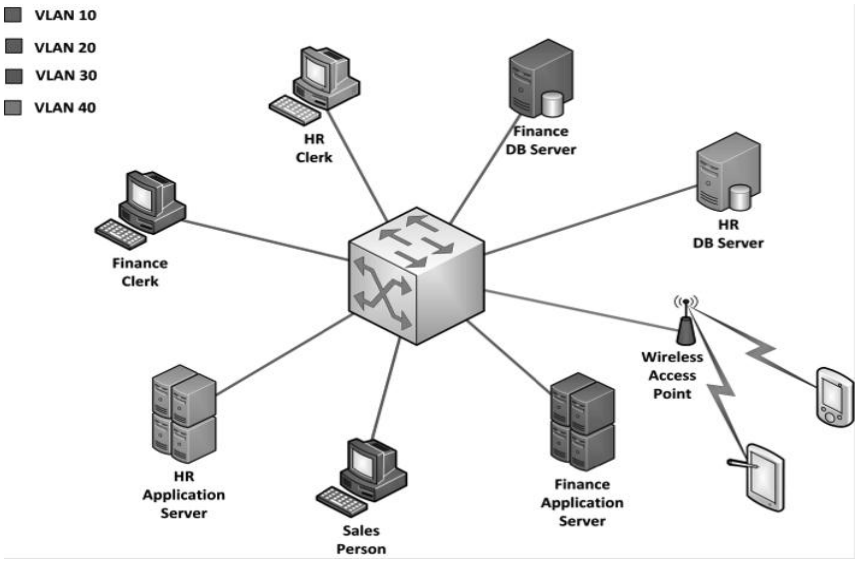



**Figure 12 :** *Exemple d'un réseau configuré en VLAN*

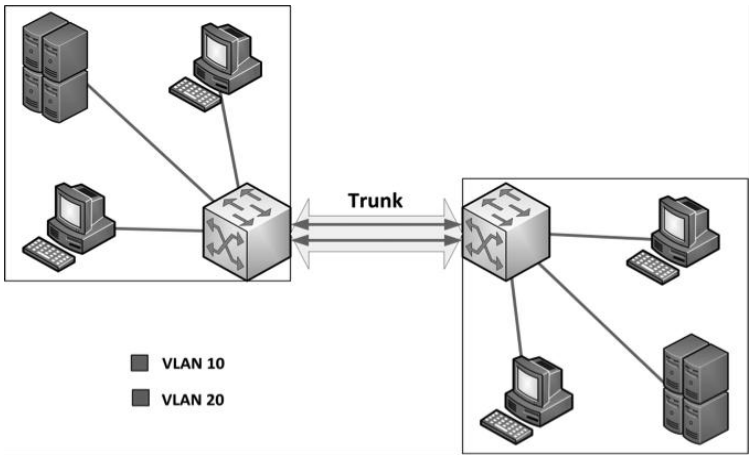



**Figure 13 :** *Exemple de configuration d'un VLAN avec trunking*

Un VLAN de niveau 1 ou VLAN par port est établi en associant les ports d'un switch à un VLAN. L'allocation des ports peut se faire de manière statique ou dynamique. Son avantage est qu'il assure l'étanchéité maximale des VLAN. Il est nécessaire d'avoir accès à des machines physiques pour pouvoir lancer une attaque. Son inconvénient est la lourdeur de l'administration. En effet, tous les switch devront être configurés manuellement un par un. Un VLAN de niveau 2 ou VLAN par adresse MAC est basé sur l'affectation des adresses MAC des équipements connectés par VLAN. Cette configuration oblige un pirate à avoir une adresse MAC autorisée pour pouvoir pénétrer au VLAN. La Table de correspondance « adresse MAC » et « VLAN » peuvent être centralisées pour éviter de configurer manuellement les switch. Ils interrogent cette table pour avoir la correspondance. En utilisant le spoofing d'adresse MAC, le pirate peut pirater le réseau. Un VLAN de niveau 3 ou VLAN par sous-réseau utilise un regroupement à partir du protocole IP. Les postes de travail vont être associés à un VLAN selon leur adresse IP. L'avantage de cette configuration est la facilité du déploiement. En effet, il suffit de configurer l'adresse du client pour qu'il joigne le VLAN voulu. Mais du point sécurité, la faille vient de cette facilité. Le pirate peut analyser les adresses IP existantes pour pouvoir pénètre à un sous-réseau. L'IP spoofing est plus facile à faire que l'écoute d'adresse MAC. Pour un réseau complexe, l'implémentation d'un VLAN peut être faite avec une architecture multi-tiers ***(Figure 14)*** en exploitant les principes cités précédents.

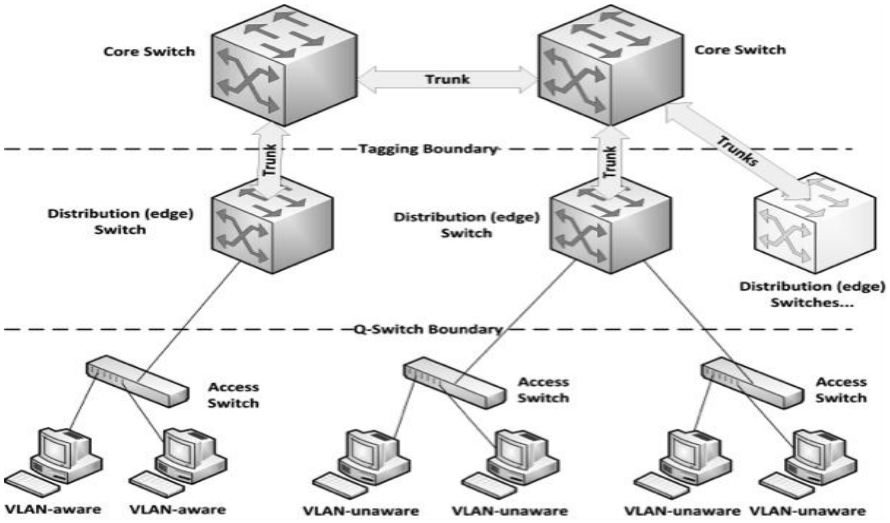



**Figure 14 :** *Exemple d'extension de VLAN*

### 4-3. Filtrage de paquets

Le filtrage des paquets consiste à mettre en place des dispositifs permettant de gérer les informations échangées. La première étape de la sécurisation est la construction d'un control d'accès et de filtrage des paquets. Le ***Tableau 1*** illustre le mode de fonctionnement d'un Q-Switch Packet Processing Les paquets sont filtrés en entrée avant d'être soumis à un analyse approfondi. Les paquets légitimes vont être soumis aux règles définies dans le « access control list ». Les paquets vont être remis au VLAN, le destinataire va encore faire des tests. La ***Figure 15*** montre le mode de fonctionnement du filtrage [12].

**Tableau 1 :** *High-level Switch VLAN Packet Processing*

| Destination Address | D-Switch | Q-Switch |
|---|---|---|
| Known MAC address | Step1: Determine the output port associated with the destination address from the CAM.<br><br>Step 2: If the associated output port is different from the port on which the packet arrived, forward the packet to the correct output port.<br><br>Step 3: Otherwise, drop the packet. | 1. Determine the VLAN associated with the packet.<br><br>2. Determine the output port associated with the destination MAC address as listed in the CAM table.<br><br>3. If the associated output port is different from the port on which the packet arrived, and it is a member of the same VLAN as the received packet, forward the packet to the output port. |
| Unknown MAC address or broadcast | Send the packet to all ports except the port on which the packet is received. | Step 1: Determine the VLAN to which the packet belongs.<br><br>Step 2: Broadcast, flood, the frame to all ports in the VLAN set except the port on which the packet is received. |

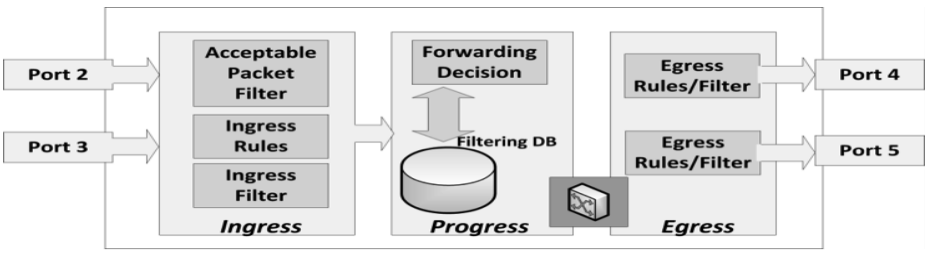



**Figure 15 :** *Q-Switch Packet Forwarding Process (Seifert & Edwards, 2008)*

L'utilisation d'une liste de contrôle réduit déjà la surface d'attaque mais reste insuffisant. La sécurisation physique est aussi nécessaire pour éviter qu'un pirate se connecte directement sur les équipements. Ensuite, il est nécessaire de protéger les accès aux différents équipements par des mots de passe et d'utiliser des protocoles sécurisés comme le SSH lors des accès à ces équipements. Les administrateurs devront avoir des rôles bien définis pour éviter les conflits de responsabilité. L'attaque de MAC Flooding, qui consiste à déborder les tables CAM des switch pour le rendre comme un HUB, est parmi l'attaque usuelle d'un VLAN. Le pirate peut utiliser par exemple l'utilitaire « dsniff » pour envoyer un grand nombre d'adresses MAC falsifiée. En faisant la résolution adresse MAC / VLAN, les switch vont être saturés. Les précautions nécessaires pour se défendre à ce type d'attaque sont l'affectation manuelle des adresses MAC à chaque port, l'utilisation de 802.1x pour forcer le filtrage de ce paquet et / ou la configuration des switch pour reconnaitre les *n* premières adresses MAC et le basculer en mode « configuration en cours ». Les ports vont rejeter les adresses MAC falsifiées pour éviter la saturation des switch. D'autres techniques existent et sont disponibles pour sécuriser les VLAN.

### 4-4. Blocage des attaques

Le blocage de l'attaque consiste à arrêter les flux non autorisés qui peuvent nuire au bon fonctionnement de l'ensemble du système. Le regroupement en sous-graphe allège déjà la complexité d'un réseau. Il offre la possibilité de définir des règles de sécurité mais reste encore insuffisant. Pour pouvoir minimiser les risques d'attaque, il faut le compléter par le principe de bloc. En effet, la méthode de base permet de décomposer un graphe en sous-composants, appelé bloc, maximales biconnexes. La ***Figure 16*** illustre un exemple de graphe à trois blocs. Tous les chemins pour aller du bloc 1 au bloc 3 ou l'inverse passe impérativement par le bloc 2. Les règles de sécurité et les autorisations vont être définies dans le bloc 2 pour assurer les communications. Ce dernier joue le rôle d'un firewall.

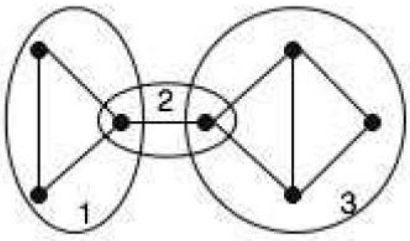



**Figure 16 :** *Graphe à trois blocs*

L'algorithme de routage glouton de Kleinberg défini le mode de transmission de message pour chaque nœud [13].

*Entrée : la source s et la cible t.*
*Initialisation : x ⟵ s. (x est le porteur courant du message)*
*tant que x ≠ t faire*
*x ⟵ y où y est le contact local ou long-distance de x qui minimise |y − t| (choix aléatoire uniforme en cas d'égalité).*
*fin tant que*

Dans ce cas, la communication entre deux nœuds de deux blocs différents doit passer par le bloc central qui effectue les filtrages de flux ainsi que les surveillances. Les firewalls et les routeurs assurent les filtrages et les aiguillages de données entre les blocs.

### *4-4-1. Firewall*

Un firewall ou pare-feu sert à contrôler les communications entre deux blocs de sous réseau. Il offre la possibilité d'implémenter des règles sur les protocoles, les équipements source et destination, et les contenus échangés. Un firewall peut être également un matériel dédié ou un logiciel. Il existe aussi des firewall qui sécurisent un seul poste de travail (non étudié dans notre cas) qui est généralement installé avec les antivirus. Au total, on a trois types de firewall : firewall sans état (stateless), firewall à état (stateful) et firewall applicatif. Un firewall sans état intervient sur les couches réseau et transport du modèle OSI. Les règles de filtrage sont basées sur l'adresse IP source et / ou destination ainsi que le protocole utilisé (port). Elles sont gardées dans un « Access Control Lists ». La menace de ce type de firewall est que pour un port ouvert par l'administrateur, les pirates pourraient exploiter aussi ce port. Ce firewall n'arrive pas à stopper les attaques de types IP Spoofing et SYN Flood. Un firewall à état est une extension de firewall sans état. Il fait en plus des fonctionnalités précédentes un control des paquets. Les attributs gardés en mémoire sont les adresses IP, les numéros de port et les numéros de séquences des paquets qui ont traversé le firewall. Il peut après autoriser ou refuser les nouvelles connexions sans lire l' « access control lists ». L'avantage est qu'il peut stopper les attaques de type DoS en supprimant les paquets falsifiées sans réponse. Ce type de protocole possède les modes de fonctionnement des protocoles standard cependant les protocoles utilisant un numéro de port spécifique demeurent limités. Le firewall applicatif fonctionne au niveau de la couche application du modèle OSI. Il prend en compte tous les protocoles utilisés par les applications. Il peut décrire une analyse plus détaillée des informations qui y transitent. Les analyses ne se basent pas tout simplement sur le numéro de port mais sur l'application. Par exemple, même si l'utilisateur change le port standard http autre que 80, le filtrage fonctionne encore. Il peut aussi utiliser comme un proxy http. Ce type de firewall consomme un beaucoup plus de ressources. Une étude de flux est nécessaire pour la réussite de la mise en place.

### *4-4-2. Gestion de flux*

Apres le regroupement et la mise en place de firewall, la dernière étape est la définition des règles d'échange de données entre les différents sous-réseaux. Le principe est de créer une matrice d'échange entre les sous-groupes. La matrice sera traduite par des règles au niveau des firewalls ou de l'élaboration de liste d'accès des VLAN. Deux éléments sont importants lors de l'établissement des règles : le protocole autorisé et le sens de la communication. La ***Figure 17*** montre un exemple simple d'un schéma de flux entre deux sous-réseaux.

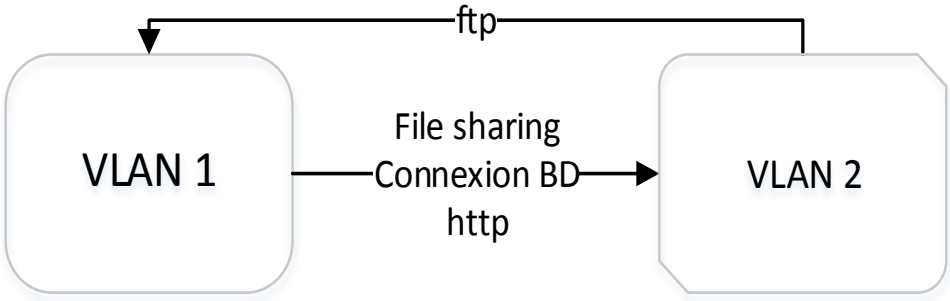



**Figure 17 :** *Exemple de schéma de flux*

L'exemple illustre que seul le ftp est autorisé du VLAN2 vers le VLAN1 et que le partage de fichier, connexion à la base de données et le http est autorisés dans l'autre sens. La représentation peut se faire aussi par une matrice par protocole.

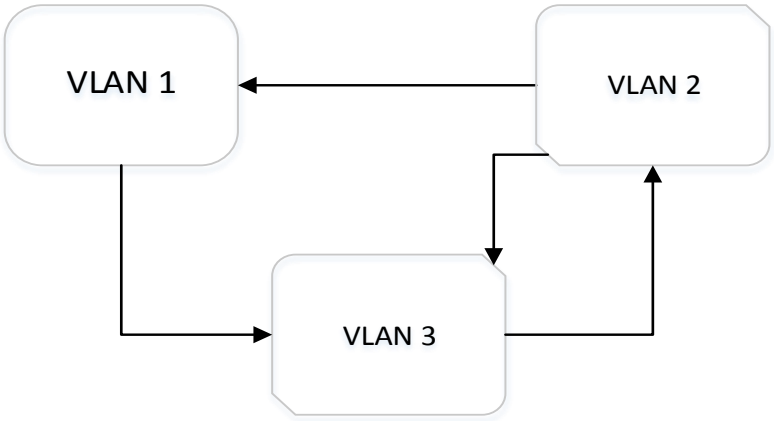



**Figure 18 :** *Exemple de schéma représentant un flux réseau pour un protocole défini*

La représentation de ce schéma ***(Figure 18)*** peut se faire par la matrice suivante $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ dont 1 symbolise l'autorisation. Dans le cas de traitement de plusieurs protocoles, la représentation peut se faire avec un ***Tableau (Tableau 2)***.

**Tableau 2 :** *Exemple de représentation de flux*

| | VLAN1 | VLAN2 | VLAN3 |
|---|---|---|---|
| VLAN1 | | ftp, http | rdp |
| VLAN2 | file sharing | | base de données |
| VLAN3 | | ftp | |

Si le regroupement est bien étudié, le schéma va être simple. Dans le cas contraire, il faut refaire l'étude de la segmentation.

### *4-4-3. Authentification matérielle et non-répudiation*

Pour assurer que l'utilisateur ne change pas d'identité et de groupe d'appartenance, il faut faire appel à l'authentification de type 802.1x. En se basant avec la théorie des graphes, ce principe consiste à fixer les nœuds. De plus, il faut aussi palier au trou de sécurité venant du GPRV des VLAN. En effet, un pirate souhaitant se connecter à un VLAN a la possibilité de configurer sa carte en respectant le standard 802.1q. Dans le cas de modélisation d'un système assurant la non-répudiation, l'authentification matérielle est obligatoire. En appliquant ce principe, les utilisateurs ne peuvent pas changer leur adresse MAC et leur adresse IP. De même aucun utilisateur ne peut changer leur poste de travail avec un ordinateur portable ou d'autre équipement comme un switch. Les informations vont être stockées dans une base de données avec les autres informations concernant chaque utilisateur. Le standard 802.1x a été mis au point par l'IEEE en juin 2001. L'objectif est de mettre en place une procédure d'authentification au moment de la connexion d'un matériel sur le réseau. Ce mécanisme d'authentification sera fait de manière transparente lors du branchement et de l'autoconfiguration comme dans le cas de DHCP. Le standard ne crée pas un nouveau protocole mais s'appuie sur les standards existants. La dialogue entre le système authentificateur et le système à authentifier se fait en utilisant le protocole EAP (Extensible Authentification Protocol) défini par le RFC2284. Les paquets sont transportés dans des trames Ethernet spécifiques EAPOL (EAP Over Lan) qui sont marquées avec le numéro de type (Ethertype) égal à 88FE (encapsulation directe d'EAP dans Ethernet). La ***Figure 19*** montre le principe [14].

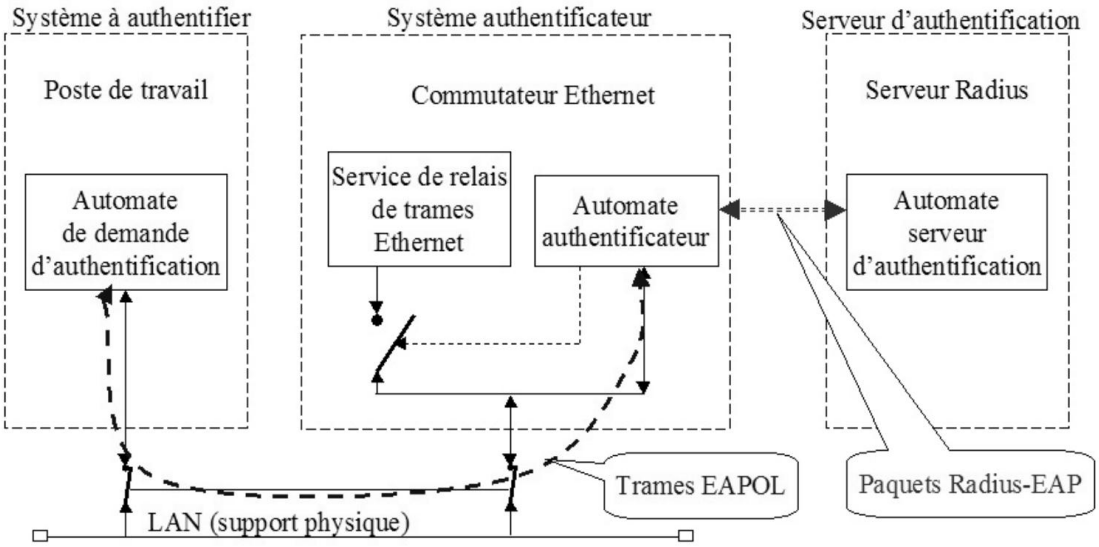



**Figure 19 :** *802.x et serveur d'authentification*

La principale faiblesse de ce mécanisme est qu'il est conçu pour la connexion physique. L'utilisation d'un hub par un utilisateur perturbera le réseau surtout s'il connait un port ouvert. L'évolution est l'utilisation d'un certificat (addendum 802.1aa).

## 5. Conclusion

« Se protéger » contre les menaces informatiques est un défi quotidien pour les détenteurs de données numériques et les fournisseurs de services. En suivant l'évolution technologique, les risques de s'exposer face à des attaques se multiplient de plus en plus. Malgré les progrès technologiques, l'homme a toujours un rôle important vis à vis de la sécurité. Dans son environnement de travail journalier, les décisions humaines peuvent avoir des conséquences importantes pour la sécurité de l'entreprise. D'après les différentes études, la première étape de la sécurisation réseau consiste à regrouper les équipements qui ont de fortes interactions et / ou les équipements critiques. La maitrise de flux qui se divise en deux parties est l'action suivante. Elle consiste à mettre en place des méthodes pour pouvoir filtré les paquets puis mettre en place des dispositifs permettant de bloquer les requêtes non autorisées. A la fin, l'authentification matérielle est nécessaire pour assurer la non-répudiation.

## Références

[1] - http://www.bbc.com/news/technology-39913630, (2017)
[2] - http://www.yzgeneration.com/wanna-cry-virus/, (2017)
[3] - http://www.lefigaro.fr/secteur/high-tech/2016/02/16/32001-20160216ARTFIG00205-un-hopital-americain-paralyse-par-des-pirates-informatiques.php, (2016)
[4] - http://www.lemondeinformatique.fr/actualites/lire-serie-d-attaques-ddos-contre-des-banques-americaines-50713.html, (2016)
[5] - https: //piratercommeunnul.wordpress.com/2014/11/02/utiliser-un-pdf-pour-pirater-votre-boss-comme-un-null/, (2014)
[6] - http://thehackernews.com/2016/04/swift-bank-hack.html, (2016)

[7] - V. LEVORATO, « Contributions à la Modélisation des Réseaux Complexes: Prétopologie et Applications », (2010)
[8] - D. MÜLLER, « Introduction à la théorie des graphes », Commission Romande de Mathématique, (2012)
[9] - P. ERDÖS et A. RENYI, « On random graphs », Publicationes Mathematicae, (1959)
[10] - https://www.aldeid.com/, (2016)
[11] - A. LAURE et T. ZUFFEREY, « Intégration de méthodes de data mining dans le renseignement criminel », (2009)
[12] - http://resources.infosecinstitute.com/vlan-network-chapter-5/, (2016)
[13] - E. LEBHAR, « Algorithmes de routage et modèles aléatoires pour les graphes petits mondes », (2006)
[14] - L. SACCAVINI, « 802.x et la sécurisation de l'accès au réseau local », (2003)