CHAPITRE 2

LES LOGICIELS MALVEILLANTS ET LES RÉSEAUX DE ZOMBIES D'OBJETS CONNECTÉS

Comme nous l'avons vu dans la partie précédente, beaucoup d'objets connectés, comme les concentrateurs, les routeurs, les caméras connectées ou les enregistreurs vidéos, utilisent un système Linux et sont vulnérables à une attaque contre les identifiants permettant de les authentifier pour utiliser l'interface d'administration. Cette vulnérabilité a été exploitée à plusieurs reprises afin de créer des réseaux de zombies. Afin de mieux comprendre les failles exploitées et ainsi, pouvoir trouver des contre-mesures adéquates, nous avons souhaité étudier les logiciels malveillants ciblant cette partie des objets connectés. Avant de décrire nos études et nos résultats, nous allons d'abord donner les définitions utiles et décrire le fonctionnement général des réseaux de zombies ainsi que les problèmes qu'ils engendrent. Nous décrirons aussi les outils les plus utilisés pour étudier ces réseaux. Enfin, nous expliciterons la problématique de recherche.

2.1 DÉFINITIONS

Nous allons ici définir les concepts de logiciels malveillants, de réseaux de zombies et nous décrirons leur organisation générale ainsi que leur cycle de vie.

2.1.1 LOGICIEL MALVEILLANT

L'équipe de Malwarebytes [1] définit un logiciel malveillant (ou « malware ») comme un terme générique qui décrit tous les programmes ou codes ayant un comportement malicieux et dommageable pour le système. Ces programmes sont hostiles et intrusifs, ils ont pour objectif de voler des données ou de stopper et d'endommager un système d'information comme un ordinateur, un réseau, une tablette ou un téléphone. Ces logiciels malveillants vont se cacher pour infecter d'autres hôtes et échapper aux systèmes de surveillance. Ces programmes peuvent aussi prendre en partie le contrôle de l'appareil infecté et donc altérer les opérations de l'utilisateur. Tout comme le virus de la grippe, les logiciels malveillants vont interférer avec le fonctionnement normal des systèmes d'information (Malwarebytes, 2019).

2.1.2 RÉSEAUX DE ZOMBIES

Les réseaux de zombies, aussi appelé « botnets » (contraction des mots « robot » et « network ») sont définis par Karspersky comme *de larges réseaux d'ordinateurs infectés par un même virus* (Fisher, 2013). Ainsi, lorsqu'un même logiciel malveillant va infecter plusieurs milliers (voir dizaines de milliers) d'objets connectés afin de les contrôler à distance, on parle de réseaux de zombies d'objets connectés. Ici, on parle de zombies du fait que les objets infectés peuvent être contrôlés en totalité par l'attaquant possédant le réseau.

Organisation et composants

Traditionnellement, on distingue deux types d'architectures pour les réseaux de zombies : centralisée ou décentralisée. La première consiste à mettre en place un serveur de contrôle et de commande (serveur C2), permettant de récupérer toutes les informations de tous les objets

1. https ://fr.malwarebytes.com/

infectés et de leur envoyer des ordres (Koroniotis *et al.*, 2019). À l'opposé, les architectures décentralisées n'ont aucun serveur central, mais utilisent un réseau pair-à-pair (« P2P ») pour communiquer. La méthode de communication centralisée a l'avantage d'être facile à implémenter et rend compte de la totalité de l'état du réseau facilement (Koroniotis *et al.*, 2019). La seconde a l'avantage d'être plus résiliente et donc plus difficile à faire tomber.

Dans le cas d'une architecture centralisée, on peut voir apparaître d'autres composants, comme le serveur de rapport ou le serveur de chargement. Ces derniers ont été introduits par Mirai (Ji *et al.*, 2018) et permettent de rendre l'architecture centralisée encore plus efficace. Le serveur de rapport permet d'enregistrer les adresses IP utilisées par des objets connectés vulnérables ainsi que leurs identifiants de connexions. Le serveur de chargement va se connecter à ces objets pour leur transmettre le logiciel malveillant. Un schéma de cette architecture est donné à la figure 2.1

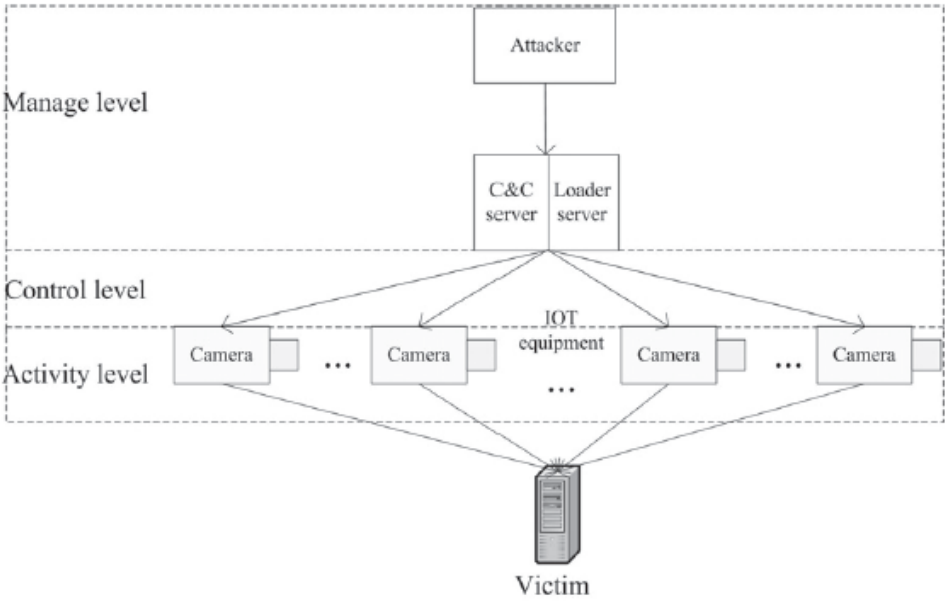

Figure 2.1 – Architecture du botnet Mirai, d'après (Ji *et al.*, 2018)

Cycle de vie

Le but principal d'un réseau de zombies est de grossir au maximum pour ensuite pouvoir lancer diverses attaques. Nous détaillerons ces dernières dans la prochaine section. Ici, nous allons décrire le fonctionnement d'un de ces réseaux, en nous basant sur le cycle de vie du programme Mirai.

Tout d'abord, il faut identifier les objets vulnérables. Pour ce faire, les réseaux de zombies doivent constamment scanner l'ensemble des adresses IP disponibles afin de trouver le maximum de victimes. Ensuite, pour chaque victime potentielle, le réseau va essayer de l'exploiter. Ensuite, si l'attaque réussit, la victime va se faire infecter et elle participera au prochain cycle. Nous pouvons résumer ce cycle en trois étapes : identification des objets, exploitation des victimes et duplication du logiciel malveillant. Un schéma récapitulatif est donné à la figure 2.2

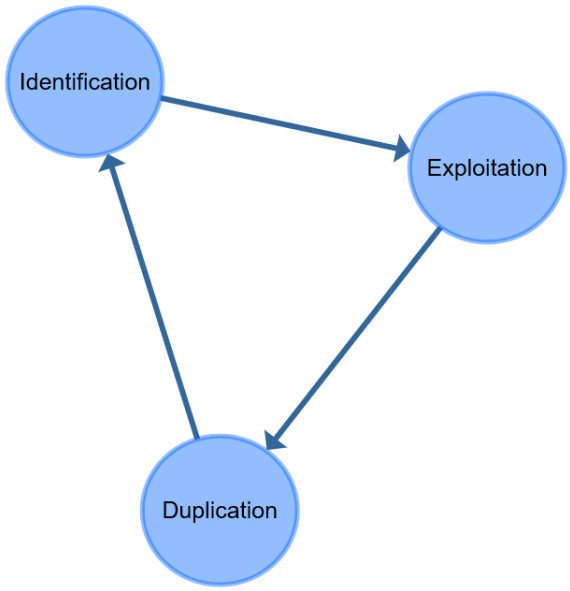

Figure 2.2 – Cycle de fonctionnement général d'un réseau de zombie grandissant

Concernant les réseaux de zombies ciblant les objets connectés, nous pouvons apporter quelques précisions sur ce fonctionnement. En effet, la phase d'exploitation est très souvent une phase

d'attaque par force brute sur les identifiants de la victime. On peut noter qu'en 2012, le réseau de zombies Carna (Carna, 2012), avait analysé l'ensemble de l'espace IP durant 24h pour créer une carte interactive des connexions au réseau Internet. Pour ce faire, il avait infecté de nombreux objets connectés qui n'avaient pas d'identifiant ou avaient des identifiants faibles. Il avait recensé plus d'un million d'objets accessibles de cette manière.

De plus, certains logiciels malveillants, comme Mirai, vont adopter une phase de dissimulation. Ils vont se charger en RAM puis supprimer leur binaire (Antonakakis *et al.*, 2017). Certains vont même changer leur nom afin d'avoir un nom de processus classique comme « telnetd » . Enfin, nous pouvons observer sur la figure 2.3 qu'une étape de rapport est mise en place. Celle-ci a été introduite par Mirai et a pour but d'améliorer la rapidité de propagation du logiciel malveillant. Cela permet à l'attaquant d'avoir une liste de tous les objets vulnérables. De plus, les zombies peuvent rapidement se remettre à scanner l'espace IP et c'est un serveur central, ayant beaucoup de puissance de calcul et de bande passante, qui s'occupe d'infecter les nouvelles victimes (Kambourakis *et al.*, 2017).

2.2 LES PROBLÈMES CAUSÉS PAR LES RÉSEAUX DE ZOMBIES

Dans cette section, nous allons détailler les principales attaques mises en place par les réseaux de zombies ainsi que leurs conséquences. D'après Jerkins (2017), les réseaux de zombies sont traditionnellement utilisés afin de créer des attaques de déni de service ou pour distribuer du « spam » (ou pourriel). Cependant, il existe aujourd'hui d'autres utilisations, par exemple, le vol d'identité, la distribution d'autres programmes malveillants ou le minage de cryptomonnaie (Koroniotis *et al.*, 2019). De manière générale, chaque réseau ne se concentre que sur un grand type d'attaque. Il existe cependant quelques exceptions, par exemple « Gameover Zeus » qui était capable de lancer des attaques de déni de services, voler des informations bancaires et lancer des campagnes de spam (Koroniotis *et al.*, 2019). Ces attaques représentent donc en

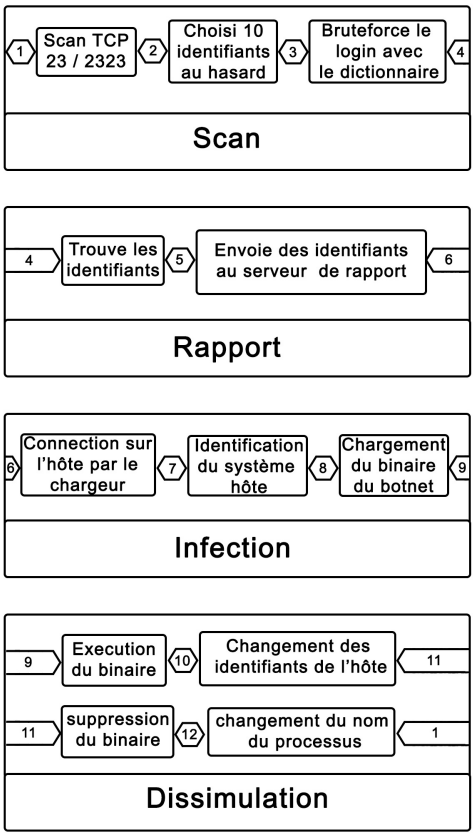

Figure 2.3 – Cycle de fonctionnement de Mirai

partie les objectifs des attaquants et en conséquences des réseaux de zombies. Nous verrons dans les chapitres 3 et 4 que ces objectifs différents peuvent amener à mettre en place des solutions différentes.

2.2.1 LE DÉNI DE SERVICE

La catégorie d'attaque la plus utilisée et la plus médiatisée est sans aucun doute l'attaque de déni de service distribuée (DDoS). Ce genre d'attaque utilise une asymétrie des ressources disponibles entre la victime et l'attaquant. Son but principal est de saturer la bande passante ou les capacités de calcul de la victime, afin que ses clients légitimes n'aient plus accès au

service. Pour mettre en place ce genre d'attaque, un grand nombre de zombies doivent utiliser au maximum leur bande passante et leur capacité de calcul afin d'envoyer le plus grand nombre possible de requêtes à la victime. On pourrait faire l'analogie avec un restaurant ou un groupe de 150 personnes viendrait occuper une salle de 100 couverts, pour demander uniquement des denrées non servies par le restaurant.

D'après Wang *et al.* (2018), le nombre et l'intensité de ces attaques étaient en croissance dans les années 2013-2014. En effet, l'équipe parle d'une augmentation de l'amplitude des attaques (taille de la bande passante utilisée) de 245% entre le premier quart de l'année 2014 et celui de 2013. À cette époque, les attaques moyennes étaient de 7.39 Gbps et la plus grosse attaque enregistrée utilisait 500 Gbps de bande passante. En 2016, les trois premières attaques du réseau Mirai étaient de 623 Gbps contre le blog de Brian Krebs et entre 1.1 et 1.2 Tbps pour les attaques contre OVH et Dyn (Antonakakis *et al.*, 2017). On voit ainsi une grande augmentation de la puissance de ces attaques au cours des dernières années.

Selon Scott Sr et Summit (2016), les réseaux de zombies sont souvent partitionnés et loués. En effet, ils montrent qu'il existe un marché appelé « DDoS-as-a-Service » qui, au lieu de proposer une application, vend des plages d'attaques de déni de service. Selon leurs travaux, le coût moyen se situe entre 25$ et 150$ par tranche de 24h contre une seule cible. Le coût varie en fonction de l'amplitude de l'attaque.

2.2.2 LE SPAM

Le concept de *spam* englobe l'ensemble des mails indésirables que l'on peut recevoir. D'après Drozhzhin (2015), ce genre de mail était à la base utilisé par les équipes commerciales, à des fins publicitaires. Cependant, les cybercriminels se sont aussi mis à utiliser les mails afin de propager leurs logiciels malveillants. En effet, c'est par le biais de pièces jointes corrompues

que se transfèrent ces programmes.

Une autre utilisation du spam par les cybercriminels est appelée hameçonnage. Cette pratique utilise l'ingénierie sociale (Kevin D.Mitnick, 2002) afin d'entraîner des usagers dans diverses arnaques. L'exemple le plus parlant est le mail d'un prince ou d'un riche homme d'affaires, vous demandant de verser de l'argent sur un compte bancaire, dans un pays différent du vôtre. Le but ici est d'utiliser les sentiments ou les émotions des victimes afin qu'elles cliquent sur la pièce jointe ou qu'elles envoient leur argent.

L'ensemble de ces techniques se base sur le fait de délivrer un nombre colossal de mails afin de toucher un maximum de personnes. Souvent, il est nécessaire qu'une petite fraction des personnes ciblées tombe dans le piège pour que l'attaquant réussisse son attaque. Par exemple, si ce dernier souhaite infecter le réseau d'une entreprise, il peut utiliser le spam afin d'envoyer une pièce jointe corrompue à l'ensemble des employés. Il n'aura besoin que d'une seule ouverture de cette pièce jointe pour infecter le réseau interne de l'entreprise. Ainsi l'utilisation de réseaux de zombies, comprenant des milliers d'ordinateurs ou d'objets connectés, facilite la distribution d'une telle quantité de courriels frauduleux.

Ici, l'avantage d'utiliser un réseau de zombies est le fait que chaque zombie enverra une fraction des mails. Ce faisant, chaque mail aura moins de chance de se faire étiqueter comme indésirable. Ce fut notamment le cas en 2014, où un réseau d'objets connectés a envoyé plus de 750 000 mails indésirables. Chaque objet n'envoyait qu'une dizaine de mails, trois fois par jour. Les chercheurs avaient même identifié un frigo parmi les zombies (Williams, 2014).

2.2.3 MINER DES CRYPTOMONNAIES

Ces dernières années, nous avons vu apparaître un engouement massif pour les cryptomonnaies. Les plus connues étant le Bitcoin (Nakamoto, 2008) et l'Etherum (Buterin, 2019). Ces dernières

sont sources de spéculation depuis quelques années, le cours du Bitcoin étant à 327 USD en novembre 2015, 17 000 USD fin décembre 2017 et à 8 600 USD au mois de novembre 2019. Le principe général de ces cryptomonnaies est basé sur le principe de « blockchain » . Ce dernier peut se voir comme un gigantesque livre de comptes publics, distribué et accessible à tous. Afin de rajouter des transactions, il faut « miner » des blocs. Le fait de miner des blocs permet de gagner des éléments de la monnaie. Ainsi, lorsqu'un mineur crée un bloc et que celui-ci est accepté par le système, il gagne des Bitcoins ou du gaz (sous division de l'Etherum). Afin de miner un bloc et donc de gagner de la monnaie, il faut effectuer une preuve de travail (« Proof of Work » ou POW). Dans le cas du Bitcoin, il faut trouver un petit nombre à ajouter au bloc, afin que le hash du bloc commence par un certain nombre de 0. Pour l'Ethereum, il faut effectuer un « contrat intelligent » , où le mineur va devoir exécuter un certain algorithme qui sera défini par le contractant.

Ainsi, l'on voit que pour obtenir ces monnaies, il faut de la puissance de calcul. Plus l'on en possède et plus nous avons de chances d'obtenir des Bitcoins ou de l'Ether. De plus, ces systèmes sont parfaitement conçus pour des systèmes distribués. Or un réseau de zombies est par définition un système distribué où les zombies travaillent en synergie pour accomplir une tâche précise. Ainsi, l'on comprend pourquoi le minage de cryptomonnaies apparaît comme des objectifs d'utilisation d'un réseau de zombies. Plus un réseau possédera de zombies et plus il sera rentable.

2.3 LES OUTILS D'ÉTUDE

Maintenant que nous avons défini les notions importantes et détaillé les objectifs principaux des réseaux de zombies, nous allons expliciter les méthodes et outils utilisés pour les étudier.

2.3.1 LES POTS DE MIEL

Les pots de miel (ou « honeypots ») sont des outils clés pour détecter et analyser les programmes malveillants (Provos *et al.*, 2004; Koroniotis *et al.*, 2019). Ce sont des objets (réel ou virtuels) qui simulent des objets faibles (ordinateur avec une faille, objet connecté sans mot de passe etc.) afin qu'ils puissent se faire infecter par un programme malveillant. Ils vont ensuite garder une copie du code pour analyse future et vont enregistrer toutes les actions que fera le programme malveillant. On distingue principalement deux catégories de pots de miels : ceux à forte interaction et ceux à faible interaction.

La première catégorie correspond à des objets simulant la totalité d'un système vulnérable et pouvant être compromis. La seconde catégorie ne simule que les services souvent attaqués et ne va pas exécuter le logiciel malveillant. L'ensemble de ces techniques permettent de facilement créer des règles de détection permettant d'identifier les futures attaques. Ces outils récoltent énormément de données sur les réseaux de zombies, au niveau de leur comportement et de leurs communications réseau. L'inconvénient majeur est le besoin de stockage qui peut rapidement exploser. De plus, il est extrêmement compliqué de simuler avec exactitude une cible adéquate. C'est encore plus vrai pour les objets connectés, qui peuvent utiliser près d'une dizaine d'architectures différentes pour les processeurs (Pa *et al.*, 2015).

Dans le monde très hétérogène de l'internet des objets, les créateurs de réseaux de zombies créent des versions différentes de leurs logiciels malveillants en fonction des architectures ciblées. On retrouve ainsi des binaires compilés pour les architectures ARM, MIPS, MIPSEL, etc. Cela oblige les chercheurs à adapter leurs pots de miels afin de capter un maximum de binaires différents et ainsi étudier au mieux les menaces. C'est ce qu'on mis en place l'équipe de Pa *et al.* (2015). Ils ont créé un pot de miel mixte, avec une partie présentant une faible interaction afin de détecter et d'enregistrer les attaques sur le protocole Telnet.

Ils ont aussi développé une partie avec une interaction forte, capable de faire tourner les programmes malveillants sur 11 architectures de processeurs différentes. Pour ce faire, ils ont utilisé de nombreuses machines virtuelles afin de simuler au mieux les objets connectés. À l'époque, cela leur a permis de découvrir 39 nouveaux échantillons, inconnus de la base de données VirusTotal. Un autre exemple est le projet SIPHON (Guarnizo *et al.*, 2017), qui met en place plus de 80 objets connectés virtuels à travers le monde. Chacun de ces objets est un pot de miel de forte interaction et utilise 7 objets réels. Leur projet n'a pas été identifié comme un pot de miel par le site Shodan, [2] montrant le réalisme de leur outil.

2.3.2 LES ANALYSE RÉSEAUX

En plus d'analyser le comportement des programmes malveillants en créant des pots de miels, les équipes de recherches analysent l'ensemble des traces réseau possibles afin de détecter les attaques mises en place par les réseaux de zombies. Pour ce faire, ils utilisent principalement trois outils : les télescopes et les scans réseau. L'ensemble de ces outils se base principalement sur des données publiques nécessaires au bon fonctionnement du réseau internet : les requêtes DNS et IP ainsi que le trafic réseau non chiffré. L'ensemble de ces outils ne peut pas analyser du trafic chiffré.

Les télescopes réseaux

Le but principal des télescopes réseau est de surveiller un grand nombre d'adresses IP (Moore *et al.*, 2004). Le télescope va enregistrer l'ensemble du trafic à destination de ces adresses. Lorsque l'on surveille une partie de l'espace IP que l'on sait non assignées, on peut facilement détecter les scans aléatoires produits par les réseaux de zombies. Cela permet aux chercheurs d'estimer en partie l'activité et la taille du réseau de zombies. De plus, en surveillant le trafic

2. `https://shodan.io`

entrant et sortant d'un grand nombre d'adresses IP assigné, il y a de fortes chances pour arriver à détecter les attaques de déni de services, la transmission de logiciel malveillant ou plus rarement les ordres d'attaques.

Ces télescopes enregistrent de très grandes quantités de données et vont rarement pouvoir traiter l'information en temps réel. Par exemple, lors de l'étude sur le programme malveillant Mirai, les équipes de Google et Akamai ont mis en place un télescope réseau enregistrant en moyenne 1,1 million de paquets par minutes, en provenance de 269 000 adresses différentes (Antonakakis *et al.*, 2017).

Ces techniques permettent en général de retracer les évènements lors d'une attaque de grande ampleur. Ce sont des outils d'enquête permettant parfois de remonter la trace des attaquants.

Les scans

La seconde méthode d'analyse réseaux couramment utilisée est le scan actif de l'espace IP. Il existe plusieurs méthodes, par exemple celle développée par Durumeric *et al.* (2015). Cela permet de rapidement récupérer des informations sur les objets vulnérables, notamment leur nature et leur fabricant (Antonakakis *et al.*, 2017). Parfois cela peut aussi donner des informations sur les systèmes d'exploitation utilisés.

Pour conclure, nous observons que les deux grandes catégories d'outils sont complémentaires. En effet, les outils d'analyse réseau permettent d'obtenir des informations sur l'épidémie que représente le réseau de zombies. Grâce à ces outils, les chercheurs peuvent obtenir des données relatives à la taille du réseau et à la nature des objets qui le composent. Les pots de miels, quant à eux, servent surtout à analyser le programme malveillant, ses caractéristiques et son comportement. Si l'on devait faire une analogie avec la médecine, les pots de miels sont les microscopes et les équipements d'analyse des agents pathogènes, là où les outils d'analyse

réseau correspondent aux outils de mesure et de gestion des pandémies.

2.4 PROBLÉMATIQUE ET QUESTION DE RECHERCHE

Nous avons vu tout au long de ces deux chapitres que le monde de l'internet des objets est complexe et hétérogène. Nous nous sommes d'abord intéressés aux objets les plus petits, ayant le moins de capacités de calcul, mais ayant un impact physique sur le monde. Nous avons vu qu'attaquer ces capteurs et effecteurs pose plusieurs difficultés, notamment le besoin de proximité. De plus, il est difficile de créer des attaques de masse en utilisant uniquement ces objets comme vecteurs de transmissions. Nous avons aussi observé qu'il est beaucoup plus simple et rentable, de créer de larges réseaux de zombies en attaquant des objets de plus haut niveau dans notre schéma architectural. En effet, une attaque sur les concentrateurs ou sur les interfaces de contrôle permet de corrompre l'intégralité d'un réseau d'objets connectés. De plus, ces interfaces étant très souvent connectées au réseau Internet, il est possible de facilement propager les logiciels malveillants afin de créer des réseaux de zombies.

Concernant les réseaux en eux mêmes, nous avons mis en évidence, au travers des outils utilisés pour les analyser, qu'ils évoluent. En effet, nous avons constaté une augmentation du nombre d'architectures de processeurs ciblées au cours des dernières années. Cela a obligé les chercheurs à faire évoluer leurs outils et leurs méthodes d'analyse. De ce fait, nous nous sommes concentrés sur l'évolution des réseaux de zombies dans le domaine des objets connectés. En effet, étant donné l'évolution constante des logiciels malveillants ciblant les objets connectés, nous avons souhaité améliorer les méthodes et techniques permettant d'analyser ce phénomène d'évolution. Notre but est de créer un outil permettant de mieux visualiser et comprendre les évolutions de tels programmes, afin de pouvoir mieux les prédire. Ainsi, la question de recherche de ce mémoire est la suivante : quelles sont les évolutions fonctionnelles des logiciels malveillants ciblant l'internet des objets et ayant pour objectif de créer un réseau de zombies ? Nous souhaitons

connaître les évolutions passées de ces programme afin de potentiellement prédire les futures évolutions.

Pour ce faire, nous avons amélioré les taxonomies déjà existantes afin de mieux décrire ces programmes, leurs objectifs et leurs méthodes d'attaques. Puis nous les avons utilisé afin de créer une nouvelle représentation de l'évolution des logiciels malveillants. Enfin, nous avons développé un outil de simulation de propagation des infections de réseaux de zombies. Cela nous permet visualiser les effets que peuvent avoir diverses fonctionnalités et comportements de programmes malveillants sur leurs vitesses de propagation et donc sur leur efficacité. Cet outil nous aide *in fine* à mieux comprendre les impacts de chaque fonctionnalité et ainsi à mieux prédire leurs évolutions.