Titre: Title:	Évaluation d'un produit de sécurité par essai clinique
Auteur: Author:	Fanny Lalonde Lévesque
Date:	2013
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Lalonde Lévesque, F. (2013). Évaluation d'un produit de sécurité par essai clinique [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1218/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/1218/
Directeurs de recherche: Advisors:	Jose Manuel Fernandez
Programme: Program:	Génie informatique

UNIVERSITÉ DE MONTRÉAL

ÉVALUATION D'UN PRODUIT DE SÉCURITÉ PAR ESSAI CLINIQUE

FANNY LALONDE LÉVESQUE
DÉPARTEMENT DE GÉNIE INFORMATIQUE ET GÉNIE LOGICIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INFORMATIQUE)
AOÛT 2013

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉVALUATION D'UN PRODUIT DE SÉCURITÉ PAR ESSAI CLINIQUE

présenté par : LALONDE LÉVESQUE Fanny
en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de :

Mme CHERIET Farida, Ph.D, présidente
M. FERNANDEZ José M., Ph.D., membre et directeur de recherche
M. MERLO Ettore, Ph.D., membre

« La vie est trop courte pour passer inaperçu. »
- Salvador Dali

À Tristan petit mousse.

REMERCIEMENTS

Les remerciements qui suivent s'adressent à tous ceux qui, de près ou de loin, ont contribué de part leur support, leur gentillesse, et leur temps à cette aventure de deux ans.

J'aimerais premièrement remercier les membres du jury pour leur temps et leurs contributions à ce mémoire. Je tiens également à remercier Trend Micro pour leur support financier, sans lequel ce projet n'aurait pu être réalisé. J'aimerais également remercier le réseau stratégique ISSNet (*Internetworked Systems Security Network*) pour leur support financier, mais surtout pour m'avoir donné l'opportunité d'échanger avec d'autres chercheurs et de rencontrer des personnes formidables.

Je tiens particulièrement à adresser mes remerciements à mon directeur de recherche, José M. Fernandez, pour m'avoir suivie et encouragée tout au long de mon projet.

Mes remerciements iront aussi à tous ceux qui ont contribué à la réussite de ce travail. Je ne peux que remercier les professeurs Sonia Chiasson et Anil Somayaji, de l'Université Carleton, pour leur temps, leurs idées, et leur support tout au long de ma maîtrise. Je tiens de plus à remercier Carlton Davis et Pier-Luc St-Onge, tout deux anciens membres du laboratoire SecSI, pour m'avoir aidé à réaliser ce projet ; ainsi que Mathieu Allard, à l'époque étudiant à la maîtrise, qui m'a aidée lors des rencontres avec les participants. Je me dois de remercier les employés de l'équipe technique ainsi que le personnel administratif du Département de Génie Informatique et de Génie Logiciel pour leur support technique et logistique. Finalement, ce projet n'aurait pu avoir lieu sans la collaboration des 50 participants qui ont bien voulu être mes cobayes.

Sur une note plus personnelle, je tiens à remercier ma famille et mes ami(e)s pour leur patience et leur support ; mon conjoint Tarek Ould Bachir, qui a su être là pour moi ; et finalement Tristan Yiddir Lalonde, mon petit mousse, pour avoir été patient toutes ces fois où il m'a accompagnée à l'université.

J'aimerais de plus profiter de ces quelques lignes pour remercier les étudiants et les membres du personnel de l'université avec qui j'ai été amenée à travailler au cours des 7 ans dernières années. Mes études à Polytechnique ont été pour moi une expérience extraordinaire, et je vous en remercie.

RÉSUMÉ

La prolifération des différentes menaces informatiques au cours des dernières années a eu pour conséquence d'augmenter la complexité des produits antivirus et de leur évaluation. Or, il est primordial de connaître la réelle efficacité des produits de sécurité afin de développer des solutions offrant une meilleure protection des systèmes informatiques face aux nouvelles menaces.

Étant donné que la majorité des tests antivirus sont réalisés dans des environnements contrôlés, plusieurs facteurs ne peuvent être pris en considération tels que l'évolution des menaces, l'interaction de l'utilisateur ou les différentes configurations du système et de l'environnement informatique. Même les tests les plus avancés, qui incluent des profils usagers automatisés, ne peuvent représenter l'ensemble des comportements usagers et des variables qui affectent l'environnement informatique.

Afin d'évaluer les produits de sécurité dans un contexte qui soit représentatif de l'expérience usager, il convient de réaliser des études holistiques, c'est-à-dire des tests où le produit de sécurité, l'ordinateur et l'utilisateur sont considérés comme un tout. Nous proposons à cet effet une nouvelle méthode d'évaluation, basée sur les essais cliniques où les participants sont invités à utiliser un ordinateur dans leur contexte habituel.

Nous présentons dans ce mémoire la première expérience du genre qui a impliqué la participation de 50 usagers durant quatre mois. Des données ont été collectées sur l'état de santé de l'ordinateur ainsi que sur le comportement des usagers afin de déterminer quel est impact de l'utilisateur sur le risque d'infection. Cette étude pilote a permis de calculer l'efficacité d'un antivirus et d'évaluer la perception qu'en ont eu les utilisateurs. Notre analyse nous a permis d'identifier des caractéristiques et des comportements usager qui exposent à un risque d'infection plus élevé. Alors que certains facteurs de risque identifiés par nos travaux peuvent sembler intuitifs, d'autres, comme le fait d'avoir un niveau élevé d'expertise en informatique, sont plus surprenants.

Le présent mémoire a permis de valider la viabilité des essais cliniques appliqués à la sécurité informatique. Qui plus est, ce travail a démontré les mérites de réaliser des essais cliniques à grande échelle afin d'évaluer les performances réelles de différents produits, techniques ou stratégies de protection des systèmes informatiques.

ABSTRACT

The ongoing growth of computer threats over the last years has increased the complexity of anti-virus products and their evaluation. However, in order to develop solutions that offer a better protection against evolving malware threats, it is important to know the true effectiveness of security products.

Since most anti-virus evaluations are performed in controlled environments, many factors (such as evolving threats, user interaction or different system configurations) can not be taken into account. Even the most advanced tests, which include automated user profiles, can not represent adequately all user behavior and other variables affecting the environment.

In order to evaluate security products in a representative way of user experience, holistic studies must be conducted. Such evaluations imply that the product to be tested, the computer and the user are considered as a whole. To that end, we propose a new evaluation method based on clinical trials, where participants are invited to use a computer in their usual context.

This thesis presents the first experiment of that kind involving 50 participants over a 4-month period. Data have been collected on the health of the computers and on users' behavior to determine risk infection factors. This pilot study allowed to evaluate the efficacy of an anti-virus product and users' perception of the product. Data analysis identified users' characteristic and behavior that increase the risk of infection. While some of the risk factors identified by our work may seem intuitive, others, such as having a high level of expertise in computers, are more surprising.

The current thesis has validated the viability of computer security clinical trials. Moreover, this work has shown the merits of conducting long-term clinical trials to evaluate the real performance of different products, techniques and strategies to protect computer system.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES ANNEXES

LISTE DES SIGLES ET ABRÉVIATIONS

URL	Uniform Resource Locator : Localisateur uniforme de ressource
HTML	Hypertext Markup Language : Language utilisé pour représenter les pages Web
BHO	Browser Helper Object : Programme additionnel qui est ajouté au navigateur Web Internet Explorer afin d'obtenir des fonctionnalités supplémentaires
NIP	Numéro d'identification personnel
DNS	Domain Name System : Système de noms de domaine
DLL	Dynamic Link Library : Format de fichier utilisé par Microsoft pour les bibliothèques logicielles. Ces bibliothèques contiennent des fonctions qui peuvent être utilisées directement par des programmes
USB	Universal Serial Bus : Bus universel en série
DVD	Digital versatile Disc : Disque numérique polyvalent
SS	Sum of squares : Somme des carrés
MS	Mean squares : Variance
MP3	MPEG-1/2 Audio Layer 3 : Format de fichier son compressé

CHAPITRE 1

INTRODUCTION

Les activités malveillantes sur Internet sont en continuelle augmentation. Selon un récent rapport de McAfee (2012), le premier quartile de 2012 a connu le plus grand nombre de programmes malveillants détectés sur ordinateurs personnels des quatre dernières années. Non seulement le nombre de menaces augmente, mais leur nature change rapidement. De simples virus parasites, les logiciels malveillants sont rapidement devenus des programmes complexes et sophistiqués développés à des fins criminelles. Les attaquants adaptent leurs techniques afin d'exploiter de nouvelles vulnérabilités, prendre avantage des nouvelles technologies et infecter plus surnoisement les ordinateurs pour éviter la détection des logiciels antivirus.

Les usagers sont plus à risque que jamais non parce que le nombre de menaces augmente, mais parce qu'ils sont de plus en plus amenés à prendre des actions directes, ou indirectes, qui entraînent l'infection de leur ordinateur. Alors que certaines de ces actions telles que l'ouverture d'une pièce jointe ou la visite d'un site Internet malveillant peuvent mener à une infection immédiate du système, d'autres, comme le fait de ne pas tenir à jour ses applications, peuvent entraîner à plus long terme une vulnérabilité du système.

Du côté défensif, les antivirus développent de meilleures méthodes de protection afin de s'adapter aux nouvelles menaces. Des premiers antivirus fonctionnant uniquement sur la détection de signatures, nous sommes passés au cours des dernières années à des logiciels complexes qui combinent plusieurs couches de protection comme l'analyse comportementale, l'identification d'URL malveillantes ou encore la classification des applications par réputation. En complément, certains antivirus vont même jusqu'à intégrer un pare-feu ou encore un système de prévention d'intrusion afin d'offrir une solution de protection complète.

L'augmentation de la complexité des menaces et des produits antivirus a eu pour conséquence de rendre plus difficile l'évaluation de ces derniers. Les méthodes typiques d'évaluation d'antivirus sont principalement basées sur des tests automatiques réalisés en laboratoire dans des environnements contrôlés. Par exemple, certains tests consistent à soumettre un échantillon de fichiers à l'antivirus ou encore à exposer ce dernier à des sites Web malveillants. Cependant, ces méthodologies ne reflètent pas les performances des produits dans un contexte réel d'utilisation puisqu'ils ne prennent pas toujours en considération plusieurs facteurs tels que l'interaction usager, l'évolution des menaces et la configuration de la machine et de l'environnement. Bien qu'il existe plusieurs méthodes d'évaluation, il n'existe aucune règle ou

norme stricte quant à leur évaluation. Par conséquent, chaque organisation ou groupe est libre d'utiliser la méthode et le protocole de son choix.

Or, la sécurité des systèmes ne peut être améliorée significativement si nous ignorons leurs faiblesses et les raisons des échecs à les sécuriser. Plusieurs hypothèses peuvent être énoncées, telles que la mauvaise programmation des applications ou la non mise à jour des systèmes. Toutefois, nul ne peut quantifier l'importance relative de ces facteurs. Qui plus est, il est impossible d'évaluer l'efficacité réelle de différents produits, techniques ou stratégies de protection des systèmes informatiques. Il est d'autant plus important de mieux comprendre l'impact du comportement utilisateur et les caractéristiques de ce dernier sur la réussite ou l'échec des mécanismes de protection.

Il est dès lors nécessaire de prendre en compte l'interaction de l'utilisateur tant avec l'antivirus qu'avec l'ordinateur. En d'autres mots, l'environnement de test devrait inclure le produit de sécurité à évaluer, l'ordinateur et l'utilisateur. De tels tests devraient permettre non seulement d'évaluer la performance des produits de sécurité, mais aussi la susceptibilité des utilisateurs à faire infecter leur ordinateur. Il devient primordial de comprendre comment les caractéristiques de l'utilisateur, telles que la perception des menaces, l'expérience en informatique ou le comportement usager, affectent le risque d'infection. Par exemple, alors que nous pourrions émettre l'hypothèse que les utilisateurs qui visitent un nombre élevé de sites Web pour adultes sont plus enclins à être infectés puisque ce type de sites est souvent utilisés par les cybercriminels pour propager des logiciels malveillants, il est important de pouvoir confirmer quantitativement une telle hypothèse.

Une nouvelle méthodologie d'évaluation basée sur les essais cliniques a été proposée par Somayaji *et al.* (2009). Cette approche implique de réaliser des études utilisateurs sur une longue période de temps. Les participants sont invités à utiliser dans leur contexte habituel leur propre ordinateur, qui a préalablement été configuré de telle sorte à déterminer l'efficacité de l'antivirus et à collecter des données sur le comportement usager. La collecte de ces données sur l'usage réel des antivirus permettrait non seulement de comprendre comment ces derniers sont utilisés mais également de déterminer quels facteurs externes influencent leur performance.

Ayant fait leurs preuves dans le domaine médical, nous sommes d'avis que les essais cliniques appliqués dans un contexte de sécurité informatique permettront de répondre à plusieurs interrogations telles que : Quels sont les bénéfices d'utiliser un logiciel antivirus sur un ordinateur personnel ? La formation des employés d'une compagnie contribue-t-elle à augmenter sa protection face aux attaques d'ingénierie sociale ? Quels sont les comportements qui exposent l'utilisateur à un risque d'infection plus élevé ?

1.1 Objectifs de recherche

Le travail qui suit vise à améliorer nos connaissances des produits antivirus afin d'augmenter leur efficacité contre les logiciels malveillants. Or, une meilleure compréhension des logiciels antivirus passe par une meilleure évaluation de leurs performances. La question de recherche adressée dans le cadre du présent projet revient à se demander comment évaluer les performances des produits de sécurité et l'importance relatives des facteurs qui influencent leurs performances. À cet effet, nous postulons que les logiciels antivirus doivent être évalués par des expériences holistiques, c'est-à-dire que l'antivirus, la machine et l'utilisateur doivent être considérés comme un tout. Une première étude pilote basée sur la méthodologie des essais cliniques a été réalisée. Plus précisement, les objectifs de recherche du projet sont :

1. Évaluer un logiciel antivirus dans un contexte réel d'utilisation, c'est-à-dire dans un environnement écologiquement réaliste avec de vrais utilisateurs. Cette évaluation sera basée à la fois sur la performance de l'antivirus en terme de détection, et sur la perception qu'en ont les utilisateurs.
2. Déterminer quel est l'impact des facteurs humains sur le risque d'infection. Plus précisement, nous tenterons d'identifier quels sont les caractéristiques et comportements qui augmentent l'exposition aux menaces, entraînant une augmentation de la probabilité d'infection. Une meilleure compréhension de ces facteurs de risque permettra de mieux orienter les efforts en terme de formation et de sensibilisation.
3. Déterminer les moyens et les sources utilisés par les fichiers malveillants afin d'infecter un système. Une meilleure connaissance pratique des mécanismes d'infection permettra entre autres de confirmer le rôle de l'utilisateur dans le processus d'infection, mais également de mieux comprendre les raisons des échecs en terme de sécurisation.
4. Valider la viabilité de la méthodologie proposée, soit le recours aux essais cliniques appliqués dans un contexte de sécurité informatique.

1.2 Travaux réalisés

Une de nos principales contributions consiste à proposer et à mettre en application une nouvelle méthode d'évaluation des produits de sécurité. Nous avons réalisé une étude utilisateurs impliquant la participation de 50 usagers durant 4 mois, ce qui constitue, au meilleur de notre connaissance, une première dans le domaine de la sécurité informatique. Une analyse préliminaire de l'antivirus et des risques d'infection a été réalisée et sera présentée. De par ses résultats, cette étude contribuera à améliorer nos connaissances des logiciels antivirus, des différents facteurs de risque d'infection et des mécanismes d'infection.

Les travaux du présent projet ont fait l'objet de trois publications. La méthodologie du projet a été présentée à la conférence *Usable Security* (USEC) (voir Lalonde Lévesque *et al.* (2012a)). Les résultats portant sur l'évaluation de l'antivirus (voir Lalonde Lévesque *et al.* (2012b)) ont été présentés dans le cadre de la 22^e conférence internationale de Virus Bulletin qui regroupe les acteurs majeurs de l'industrie antivirus. Un troisième article (voir Lalonde Lévesque *et al.* (2013)) portant sur l'impact des facteurs humains sur le risque d'infection a été accepté à la conférence *ACM Computer and Communications Security* (CCS).

1.3 Plan du mémoire

Le reste du mémoire est divisé en 5 chapitres. Le chapitre 2 présente des notions de base relatives aux logiciels malveillants et aux méthodes d'infection, en plus d'expliquer le fonctionnement des logiciels antivirus et les méthodes d'évaluation actuellement utilisées. Le chapitre 3 décrit le processus de certification du protocole ainsi que la méthodologie expérimentale utilisée pour réaliser l'essai clinique. Le chapitre 4 présente une analyse et discussion des résultats obtenus quant à l'évaluation de l'antivirus et à l'impact des facteurs humain sur le risque d'infection. Le chapitre 5 explique les limitations et les améliorations possibles de la présente expérience, en plus d'exposer comment la méthode utilisée peut être adaptée afin de réaliser des évaluations comparatives entre plusieurs produits antivirus. Finalement, le chapitre 6 présente les conclusions du présent projet de recherche.

CHAPITRE 2

LOGICIELS MALVEILLANTS ET ANTIVIRUS

Ce chapitre est une introduction aux différents concepts jugés nécessaires à la bonne compréhension du présent projet. Nous y présentons des notions de base relatives aux programmes malveillants, au fonctionnement des logiciels antivirus, ainsi qu'une revue critique des différentes méthodes d'évaluation d'antivirus.

2.1 Logiciels malveillants

Un logiciel est considéré comme étant malveillant s'il a été conçu dans le but de détruire, endommager ou détourner l'utilisation légitime d'un système informatique et ce, sans le consentement de l'utilisateur. À titre d'exemple, nous entendons par logiciel malveillant, ou maliciel, les virus, les vers, les chevaux de Troie et autres menaces. Bien qu'il existe plusieurs catégories de logiciels malveillants, seulement les plus répandues sont présentées.

2.1.1 Classification des menaces

Virus informatique

Considérés comme la plus vieille famille de logiciels malveillants, les virus informatiques tirent leur nom d'une analogie avec les virus en biologie. Tout comme leur pendant biologique, les virus informatiques se reproduisent en infectant un logiciel légitime ou un document qui leur sert d'hôte.

La famille des virus peut elle-même être divisée en plusieurs catégories selon la fonction principale et la méthode de propagation utilisée (Hansman et Hunt (2003)). Parmi les types les plus courants, le virus classique, aussi nommé virus infecteur de fichier ou virus parasite, est un morceau de programme qui s'intègre dans un exécutable légitime. Plus spécifiques, les macrovirus sont des programmes développés dans le language utilisé par l'application visée. Toutefois, ce type de virus ne vise pas à infecter l'application hôte mais plutôt les documents crées en y ajoutant des macrocommandes. D'autres virus, comme les virus de secteur d'amorçage, s'attachent au secteur de démarrage d'un support amovible (disquette, disque dur, clef USB, cédérom) afin d'être exécuté automatiquement.

Ver informatique

Contrairement aux virus informatiques qui nécessitent la présence d'un hôte pour se reproduire, les vers informatiques, *computer worms* en anglais, sont des programmes autonomes. Afin d'assurer leur reproduction, ils exploitent les failles de sécurité pour se propager dans ordinateur à un ordinateur.

Les vers informatiques peuvent être divisés en deux catégories principales selon leur méthode de propagation (Hansman et Hunt (2003)). La première catégorie repgroupe les vers d'envoi de courrier en masse. Comme leur nom l'indique, ces vers ont la particularité de se propager par l'intermédiaire du courrier électronique en utilisant le plus souvent le carnet d'adresses d'un client de messagerie électronique. La deuxième catégorie contient les vers dits sensibles aux réseaux, c'est-à-dire qu'ils analysent les réseaux afin d'identifier des hôtes vulnérables.

Cheval de Troie

Nommés en référence à la l'épopée d'Illiade de Homère, les chevaux de Troie, *Trojan horses* en anglais, sont des programmes d'apparence légitime qui exécutent des actions malveillantes à l'insu de l'utilisateur. Contrairement aux virus et aux vers informatique, les chevaux de Troie ne cherchent pas à se propager mais à introduire un ou plusieurs programmes malveillants sur le système infecté.

Porte dérobée

Une porte dérobée, *backdoor* en anglais, est un maliciel qui permet de contourner le processus d'authentification régulier d'un système afin d'y gagner accès. Souvent installée par un attaquant ayant gagné accès au système, elle permet d'obtenir un accès distant non autorisé à un système tout en s'assurant de ne pas être détecté et ce, même si les failles de sécurité présentes sur le système sont corrigées.

Logiciel publicitaire

Un logiciel publicitaire ou encore publiciel, *adware* en anglais, est un programme qui affiche de la publicité lors de son utilisation ou encore lors de son installation. Bien qu'à la base ces programmes soient légitimes, certains d'entre eux contiennent du code malveillant, ce qui en fait des logiciels dits espions.

Logiciel espion

Les logiciels espions, *spyware* en anglais, sont des logiciels malveillants qui ont comme principale fonction de collecter et d'envoyer des informations sur l'utilisation du système. Souvent inclus lors du téléchargement de logiciels gratuits, ils sont pour la grande majorité installés à l'insu de l'utilisateur. Tout comme les logiciels publicitaires, les logiciels espions peuvent être utilisés à des fins publicitaires. Ils sont souvent développés par des sociétés proposant de la publicité sur Internet et permettent à ces dernières d'envoyer de la publicité ciblée selon le comportement de l'utilisateur.

Logiciel de sécurité falsifié

Un logiciel de sécurité falsifié, *rogueware* en anglais, est un programme qui se fait passer pour un logiciel de protection tel un antivirus. Ces programmes cherchent à convaincre les utilisateurs que leur ordinateur est infecté afin de leur vendre un faux logiciel. L'infection peut être fictive, dans un tel cas le faux logiciel acheté est inutile, ou lui-même malveillant.

2.1.2 Évolution des logiciels malveillants

De la théorie à la pratique

L'idée de développer des programmes capables de se reproduire fut introduite pour la première fois en 1949 par Von Neumann avec sa théorie des automates auto-reproducteurs (Von Neumann (1951)). Initialement développés comme des applications innofensives, les premiers logiciels malveillants qui sont apparus dans les années 1970 visaient soit à jouer un tour ou encore à démontrer des principes d'intelligence artificielle. Ils étaient davantage destinés à déranger les utilisateurs et à faire connaître leurs créateurs. À titre d'exemple, Creeper, le premier ver Internet, était innofensif pour l'usager ainsi que pour l'ordinateur. Il se contentait d'infecter les ordinateurs et d'afficher le message : "I'M THE CREEPER, CATCH ME IF YOU CAN". En réponse, un programme nommé Reaper fut crée afin de se propager et de supprimer Creeper des ordinateurs infectés, ce qui lui valu d'être considéré comme un des premiers logiciels antivirus. Ces applications pseudo-malicieuses étaient faciles à supprimer et ne représentaient aucune menace pour la sécurité des systèmes infectés. Elles utilisaient comme principal vecteur d'infection des dispositifs de stockage externes tels que les disquettes, les lecteurs amovibles, et les cédéroms, rendant leur propagation très lente en comparaison des logiciels malveillants d'aujourd'hui. Les créateurs ne se préoccupaient pas de dissimuler leurs programmes, essayant au contraire d'attirer l'attention sur leurs réalisations.

À la conquête d'Internet

Les choses ont drastiquement évoluées avec l'arrivée d'Internet. Alors que la propagation des premiers virus dépendait principalement de l'usager, de nouvelles menaces ont commencé à exploiter des failles dans le code ou la logique des applications afin de pénétrer un système et ce, sans qu'aucune action ne soit requise par l'usager. De l'époque où les développeurs de logiciels malveillants cherchaient la gloire, nous sommes passés à des développeurs davantage intéressés à infecter un plus grand nombre d'ordinateurs le plus rapidement possible. Ce changement de stratégie donna lieu en 1999 à la première épidémie mondiale causée par le virus Melissa. Ce macrovirus avait la particularité de se propager massivement par courriel suite à l'ouverture d'un fichier Word infecté prétendant donner des mots de passe pour accéder à des sites pornographiques.

Avec le temps, les attaquants ont adapté leurs méthodes d'infection afin de convaincre l'usager d'exécuter un fichier infecté. Le ver ILOVEYOU lancé en 2000 est un des premiers exemples notables à utiliser l'ingénierie sociale, qui consiste à manipuler les personnes afin de contourner les dispositifs de sécurité. Plus présicement, ce ver envoyait une lettre d'amour sous la forme d'un fichier texte qui contenant en fait du code *Visual Basic* malveillant.

Un nouveau modèle d'affaires

Jadis utilisés comme outils de vandalisme électronique, l'écriture et la propagation des logiciels malveillants sont aujourd'hui un marché important valant plusieurs millards de dollars par année. Les virus tels que nous les connaissions ont progressivement été remplacés par des chevaux de Troie, des vers et des réseaux de bots conçus pour générer des revenus.

Les logiciels malveillants permettent aujourd'hui d'infecter un système afin de réaliser un objectif précis comme l'envoi de spam, le vol de numéros de cartes de crédit, l'affichage de publicités ou encore l'ouverture d'une porte dérobée dans le réseau d'une entreprise. Pour parvenir à leurs fins, les cybercriminels ont raffiné leurs méthodes de propagation en impliquant à nouveau l'usager dans le processus d'infection. Une méthode de plus en plus répandue nommée le *Drive-by download* consiste à infecter discrètement un ordinateur par un programme fraudulaux. L'infection peut résulter d'une simple visite d'un site Web, de l'ouverture d'un courriel HTML ou encore du téléchargement d'un programme dont les vraies intentions sont inconnues de l'usager.

Selon un rapport de Microsoft sur les méthodes de propagation, 45% des infections requièrent une intervention de l'utilisateur, 43% proviennent de dispositifs de stockage externe, 6% sont dues à des exploits accessibles à distance et 6% proviennent soit d'une macro *Microsoft Office*, d'une attaque par force brute ou encore d'un virus infecteur de fichier (Faulhaber

et Felstead (2011)). La combinaison des deux premières méthodes de propagation permet de constater que le succès de la quasi-totalité (88%) des infections reposent aujourd'hui sur l'usager.

2.2 Logiciels antivirus

Les logiciels antivirus sont la première ligne de défense contre les logiciels malveillants. Ces programmes informatiques ont pour fonctions d'identifier, d'arrêter et de supprimer les logiciels malveillants présents sur différents types de stockage.

2.2.1 Industrie des antivirus

Le nombre de logiciels malveillants a augmenté de manière exponentielle au cours des dernières décennies. Des premiers maliciels apparus dans les années 1980, à quelques centaines dans les années 1990, il existe aujourd'hui plusieurs millions de variantes de logiciels malveillants. Cette prolifération des maliciels a eu pour conséquence d'augmenter la croissance de l'industrie des produits antivirus. Uniquement au cours de la dernière année, le marché aurait augmenté de 7,9% en 2012, passant de 17,7 à 19,2 milliards de dollars US selon le dernier rapport de la firme Gartner, spécialisée dans la recherche en technologie de l'information (Contu et Cheung (2013)).

Afin de contrer les nouvelles menaces, les usagers ont de plus en plus recours à différentes solutions afin de protéger leurs systèmes. Selon un étude réalisée par McAfee auprès de plusieurs millions d'usagers (McAfee (2012)), 83% des ordinateurs personnels sont protégés par un produit de sécurité, tel qu'un logiciel antivirus fonctionnel, un logiciel anti-espion ou un pare-feu. En fait, l'utilisation des logiciels antivirus est maintenant considéré comme la solution de protection la plus utilisée (AV Comparatives (2013b)).

Il existe actuellement plusieurs dizaines de produits antivirus disponibles sur le marché. Parmi les principaux joueurs de l'industrie, nous retrouvons entre autres les compagnies Symantec, McAfee et Trend Micro, qui occupent à elles seules plus du tiers du marché. Alors qu'une majorité de compagnies offrent des produits payants, d'autres, comme AVG, Avast et Antivir, offrent des solutions gratuites. Bien que les antivirus payants dominent le marché, la part des logiciels gratuits serait en augmentation depuis les dernières années (AV Comparatives (2013b)).

2.2.2 Méthodes de détection

Afin de détecter les fichiers potentiellement malveillants, l'antivirus a recours principalement à trois méthodes de détections : l'analyse par signatures, l'analyse heuristique et l'analyse comportementale.

Analyse par signatures

Cette méthode de détection consiste à analyser les fichiers présents sur un système afin de détecter la présence de signatures associées à des fichiers malveillants connus. Par signature, nous entendons une partie du code, ou la totalité, qui permettrait l'identification du logiciel malveillant. Ce principe de détection repose sur l'utilisation de la base de données de signatures de l'antivirus, laquelle doit être mise à jour régulièrement afin de garantir la protection du système.

Analyse heuristique

Le mot heuristique provient du mot Grec « heuriskein », qui signifie découvrir. L'analyse heuristique décrit une méthode qui permet de découvrir d'éventuels logiciels malveillants qui ne sont pas encore connus de l'antivirus. Son fonctionnement consiste à analyser le comportement supposé des programmes afin de déterminer si ces derniers sont malveillants ou non.

Analyse comportementale

Cette technique de détection repose sur l'analyse permanente du comportement des logiciels actifs sur le système. Elle permet de détecter et de bloquer l'action d'un programme considéré comme potentiellement malveillant et de prévenir les dommages sur le système visé.

2.2.3 Évolution des logiciels antivirus

Depuis les 15 dernières années, les logiciels antivirus ont évolué de concert avec les logiciels malveillants afin de s'adapter aux nouvelles menaces. Des premières solutions antivirus à aujourd'hui, quatre générations de produits peuvent être définies (Bashari Rad *et al.* (2011), Szor (2005)).

Première génération

La première génération de logiciels antivirus était uniquement basée sur la détection de signatures afin d'identifier des virus. Le logiciel analysait des fichiers afin de détecter une ou

plusieurs séquences d'octets associées à des virus connus. Or, cette méthode n'est idéale que si les séquences choisies sont propres aux virus et qu'elles n'existent pas dans des fichiers légitimes.

Un des désavantages de ce type d'antivirus est qu'ils demandent une mise à jour régulière de la base de données. De plus, ce type de détection n'est efficace que dans le cas où le virus est connu de la base de données, ce qui limite la protection de l'utilisateur face à de nouveaux virus (Helenius (2002)).

Deuxième génération

La seconde génération d'antivirus s'est dévelopée au début des années 2000 lorsque les premières générations commencèrent à perdre de leur efficacité face aux nouveaux virus. Contrairement à la génération précédente, les antivirus de deuxième génération se basent sur des identifications quasi-complètes, voir complètes, ce qui rafine leur détection des virus.

Ces nouveaux types d'antivirus introduisent la notion de détection heuristique, ce qui leur permet d'identifier de nouveaux fichiers malveillants. Cette nouvelle technique d'analyse peut être classifiée en deux catégories, statique ou dynamique. La version statique consiste à analyser la structure et l'organisation du code, alors que la version dynamique émule le code du virus afin d'y détecter des comportements malveillants (Bashari Rad *et al.* (2011)). Toutefois, le principal inconvénient de cette méthode est le nombre de faux positif générés, c'est-à-dire le nombre d'applications légitimes qui sont classifiées comme étant malveillantes.

Troisième génération

Cette nouvelle génération se distingue de ses prédecesseurs en analysant le comportement des fichiers au lieu de leur code ou de leur structure. Cette technique présente l'avantage de ne pas faire intervenir de signatures ni de règles heuristiques, tout en garantissant la protection des usagers face à de nouvelles menaces.

Le principal inconvénient de cette méthode est qu'il est parfois difficile de faire la différence entre une action malveillante et une action légitime. Par conséquent, les antivirus continuent d'être principalement basés sur la détection de programmes malveillants connus.

Quatrième génération

De nos jours, la majorité des solutions antivirus peuvent être classifiées de quatrième génération. Ces nouveaux logiciels antivirus combinent une variété de méthodes de détection telles que la détection par signatures et par règles heuristiques et l'analyse comportementale. Certains produits vont en plus intégrer des techniques de protection additionnelles comme le

recours à un pare-feu, soit un système qui contrôle le trafic réseau selon des règles définies, ou encore un système de prévention d'intrusion qui analyse les activités réseau et système afin de détecter la présence d'activités malveillantes.

2.3 Évaluation des logiciels antivirus

L'évaluation des produits antivirus est devenue une composante vitale afin d'aider l'industrie à développer de meilleurs produits. (Kosinar *et al.* (2010)). Qui plus est, les usagers attendent des logiciels antivirus que ces derniers leur fournissent une protection complète. Il devient dès lors nécessaire de réaliser des tests d'antivirus compréhensibles et réalistes, qui sauront guider les usagers afin de choisir un logiciel qui réponde à leurs besoins (Morales *et al.* (2010)).

2.3.1 Marché des tests antivirus

L'augmentation continuelle du nombre de solutions antivirus disponibles sur le marché a eu pour conséquence d'accroître la compétition entre les différentes compagnies. Par conséquent, les grands joueurs de l'industrie antivirus sont prêts à investir d'importantes sommes d'argent dans la réalisation de tests compte tenu des enjeux que peuvent représenter les résultats.

Parmi les types de tests les plus fréquents, nous retrouvons entre autres les tests individuels réalisés par les compagnies antivirus ou encore les tests indépendants sponsorisés par des compagnies antivirus. Un des principaux problèmes avec ce type de tests est qu'ils sont souvent limités à une description qui présente le produit à son avantage. De plus, lorsqu'une expérience est réalisée, cette dernière est souvent exécutée à partir d'un échantillon de logiciels malveillants fournis par la compagnie antivirus elle-même (ESET (2009)).

Il existe cependant plusieurs organisations indépendantes qui se spécialisent dans l'évaluation des logiciels antivirus. Parmi les plus reconnues, nous retrouvons entre autres AV Test, NSS Labs, AV Comparatives, Dennis Technology Labs, PC Security Labs et Virus Bulletin. Ces différentes organisations offrent une variété de tests allant du test de certification au test comparatif d'antivirus.

Alors que certains tests visent à évaluer les fonctionnalités globales des produits antivirus, d'autres se spécialisent sur des aspects précis tels que le taux de détection, le taux de faux positifs, l'utilisabilité, la désinfection ou encore la performance. Ce dernier type de test vise à évaluer plusieurs aspects du produit antivirus comme l'utilisation de la mémoire, le temps de démarrage à l'ouverture de l'ordinateur, la vitesse d'analyse de fichiers, etc.

2.3.2 Normes et éthique des tests antivirus

Les tests de logiciels antivirus ne sont encadrés par aucune norme ou règlementation officielle. Toutefois, plusieurs efforts ont été faits en ce sens afin de fournir des lignes directrices. Différentes organisations telles que EICAR (*European Institute for Computer Antivirus Research*) et AVIEN (*Anti-virus Information Exchange Network*) ont développé leurs propres codes de conduite en ce qui concerne les tests de logiciels antivirus.

En 2008, l'organisation AMTSO (*Anti-Malware Testing Standards Organization*) a été créee afin d'améliorer la qualité, la pertinence et l'objectivité des tests d'antivirus. Regroupant plusieurs acteurs tant du milieu académique qu'industriel, l'organisation vise entre autres à fournir des principes de base aux testeurs de logiciels antivirus (voir Anti-Malware Testing Standards Organization (2008b)).

Parmi ces principes, nous retrouvons que les tests réalisés ne doivent pas mettre en danger le public. En d'autres mots, les testeurs doivent s'assurer de ne pas propager les logiciels malveillants utilisés dans le cadre de leurs tests et de ne pas créer de nouveaux logiciels malveillants. Autre aspect important, AMTSO encourage les testeurs à fournir une description de la méthodologie utilisée afin de garantir une meilleure transparence des tests.

Bien qu'il n'existe aucune certification au niveau des testeurs ou de la méthodologie de test, la majorité de la communauté antivirus s'entend sur certains principes de base comme ne pas créer de nouveaux logiciels malveillants, en acheter ou en vendre, et prendre les précautions nécessaires afin d'empêcher la propagation de ces derniers (Helenius (2002)).

2.3.3 Méthodes de test

Test statique

Cette technique consiste à tester l'efficacité de l'antivirus en lui soumettant un échantillon de fichiers potentiellement malveillants qui ne seront pas exécutés. De par sa simplicité, ce type de test peut être réalisé simplement en lançant une analyse sur demande de l'antivirus. De plus, cette méthode d'évaluation n'est pas dépendante de la plate-forme utilisée.

Cependant, compte tenu du fait qu'il n'y a pas d'exécution de fichier, et donc aucun comportement à observer, ce type de test peut désavantager les produits antivirus qui utilisent des techniques dites actives et proactives de détection (ESET (2009)). Par conséquent, les résultats obtenus ne peuvent refléter adéquatement la capacité de détection des produits testés.

Test dynamique

À l'opposé des tests statiques, les tests dynamiques impliquent l'exécution des fichiers à tester, de sorte à exposer en temps réel l'antivirus à différentes menaces (Anti-Malware Testing Standards Organization (2008a)).

Bien que les tests dynamiques soient meilleurs en principe, ils présentent plusieurs inconvénients. Ces tests sont souvent difficiles à définir et à réaliser en plus d'avoir des coûts plus élevés et d'être plus demandant en terme de ressources (ESET (2009)).

2.3.4 Environnement de test

Échantillon de menaces

Un des principaux problèmes avec les tests réalisés en laboratoire se pose au niveau du choix de l'échantillon des maliciels (Gordon et Ford (1996)). Souvent, cet échantillon est soit trop petit ou aucunement représentatif de la réalité puisqu'il contient des logiciels malveillants « fabriqués » pour les tests, ou encore de vrais logiciels malveillants qui ne représentent plus les tendances observées (Kosinar *et al.* (2010)). Un autre facteur important à prendre en considération est la notion du temps. La création d'une banque de maliciels est une longue opération alors que de nouveaux types de virus sont crées tous les jours (Halloran *et al.* (1999)).

L'organisation *The WildList Organization International* a proposé une collection de virus fournis par la communauté. Cette liste a comme principal avantage d'avoir été validée par des experts, ce qui réduit le risque de faux positif. Elle contient uniquement des virus qui ont été observés à « l'état sauvage ». De plus, elle diminue le risque de certains biais, comme le biais géographique, puisque tous, indépendamment de leur localisation, peuvent y contribuer (Harley (2009)). Cependant, cette liste n'est pas nécessairement représentative de la totalité des logiciels malveillants. De plus, elle n'est mise à jour que mensuellement, ce qui donne le temps aux compagnies d'antivirus de détecter et d'intégrer ces mêmes virus à leur base de données (Clementi (2007)). Il devient donc presque impossible de réaliser des tests avec un échantillon de maliciels qui représentent les conditions réelles.

Depuis quelques années, certaines organisations commencent à offrir des tests dits en temps réel (Dunlap (2009)). Ces tests consistent soit à exécuter des fichiers ou encore à exposer l'antivirus à des sites Web compromis. À titre d'exemple, le NSS Labs a lancé en 2009 une nouvelle méthodologie de tests où les antivirus sont exposés à plus de 350 nouveaux logiciels malveillants par jour durant la période d'évaluation. D'autres organisations telles que Westcoast Labs, AV-Comparatives et Dennis Technology Labs offrent aussi des tests en temps réel. Cependant, la grande majorité de ces compagnies ne divulguent pas la méthode

utilisée quant à la sélection des menaces afin de préserver la confidentialité de la méthodologie ou encore, afin d'éviter que les compagnies antivirus ne tentent de prédire la composition de l'échantillon. Par conséquent, il est presque impossible d'en évaluer le biais.

Bien que ces tests soint plus réalistes au niveau des logiciels malveillants, ils ne peuvent réfléter les performances réelles des produits puisqu'ils font, entre autres, abstraction de l'utilisateur. Il devient nécessaire de répliquer certains comportements comme par exemple la visite de sites Web, le téléchargement de fichiers, la simulation d'attaques basée sur l'ingénierie sociale, l'exploitation de vulnérabilités, etc.

Interaction de l'utilisateur

Il est toutefois difficile de prendre en compte l'interaction de l'utilisateur lors de tests réalisés en laboratoire. Certaines organisations tentent d'adresser ce problème en utilisant des applications qui simulent la souris, le clavier et l'interaction avec de vrais programmes sur l'ordinateur (Vrabec et Harley (2010)). Toutefois, le problème majeur de ces tests est que chaque utilisateur est différent et qu'il est impossible de réaliser un scénario de tests par défaut qui sera représentatif (Kosinar *et al.* (2010)).

Vrabec et Harley (2010) et Halloran *et al.* (1999) ont proposé comme alternative de créer différents scénarios de tests adaptés selon certains profils utilisateurs. À titre d'exemples, un utilisateur ayant un profil d'internaute peut être simulé par un script qui visite plusieurs sites Web alors qu'un utilisateur ayant un profil de joueur peut visiter différents sites de jeux en ligne et télécharger des jeux. Non seulement l'utilisateur peut être simulé, mais le type de tests peut aussi être adapté selon le profil. Par exemple, les tests réalisés pour un utilisateur présentant un profil de joueur devraient être orientés vers la latence du réseau ou encore la dégradation de l'affichage par seconde, alors que les tests pour un utilisateur de type travailleur devraient par exemple mettre l'emphase sur le téléchargement de fichiers à partir d'un serveur ou encore l'édition de fichiers vidéo et audio. Le fait d'évaluer les antivirus sous différents profils devrait permettre d'exposer les forces et les avantages de chaque produit dans un contexte se rapprochant d'une utilisation réelle et ce, selon le profil de l'usager.

Une première expérience a été réalisée en ce sens en 2013 par PC Security Labs (2013). Regroupant au total sept profils d'utilisateurs, l'expérience a tenté d'évaluer les performances de différentes solutions antivirus en prenant en compte les besoins spécifiques de chaque type d'utilisateur en simulant leurs comportements par des scripts. De par ses résultats, cette expérience a démontré que les performances d'un produit antivirus peuvent varier en fonction du profil de l'utilisateur. Cependant, nous ne savons pas jusqu'à quel point ces scénarios simulent de façon réaliste l'expérience réelle des utilisateurs.

2.3.5 Essai clinique d'antivirus

Les tests réalisés en laboratoire sont essentiels mais ils ne peuvent refléter les performances des produits dans un contexte réel d'utilisation. Ils ne prennent pas toujours en compte l'interaction de l'antivirus avec plusieurs facteurs tels que le système, les applications installées, les usagers et les différents types d'attaques. Il devient dès lors nécessaire d'évaluer les logiciels antivirus par des expériences dites holistiques, c'est-à-dire dans un environnement où le système, l'antivirus et l'usager sont considérés comme un tout.

Comme alternative, Somayaji *et al.* (2009) ont proposé de s'inspirer de la méthodologie des essais cliniques utilisée dans le domaine médical. Cette méthode de recherche permet entre autres d'évaluer l'efficacité et la tolérance d'une méthode diagnostique ou d'un traitement. Appliquée dans un contexte de sécurité informatique, cette méthode implique de réaliser des études utilisateurs afin d'évaluer l'efficacité d'un antivirus à protéger un système contre différents logiciels malveillants.

Le recours à de telles expériences est particulièrement fréquent dans la discipline de l'utilisabilité, qui consiste à prendre en compte l'aspect humain dans l'étude de la sécurité des systèmes. Ces études peuvent prendre différentes formes, allant des expériences réalisées dans un environnement contrôlé aux études de terrain où le système est déployé et étudié dans son contexte réel d'utilisation. Ce dernier type d'expérience est particulièrement d'intérêt puisqu'il permet d'observer dans des conditions réelles comment les usagers interagissent avec les systèmes informatiques. Rode (2009) a examiné le comportement parental quant à la protection en ligne des enfants, Wash (2010) a eu recours à des questionnaires pour comprendre le modèle mental des utilisateurs à l'égard de la sécurité informatique et De Luca *et al.* (2010) a conduit une expérience observationnelle sur les guichets automatiques afin d'étudier l'utilisation des numéros d'identification personnel (NIP). Toutefois, le recours aux expériences dites de terrain reste encore peu commun, principalement en raison des coûts, du temps demandé et des risques potentiels de sécurité pour les usagers. Par conséquent, la majorité des études d'utilisabilité sont réalisées en laboratoire avec un nombre restreint d'utilisateurs qui dépasse rarement 50.

Alors que certaines études mentionnent l'utilisation de produits antivirus comme mesure de protection employée par les utilisateurs, aucune d'entre elles n'est orientée vers l'évaluation des antivirus. Au meilleur de nos connaissances, il n'existe à ce jour aucune étude publiée visant l'évaluation de solutions antivirus avec de vrai usagers.

2.4 Conclusion

Nous avons vu dans ce chapitre comment la réalité des logiciels malveillants a rapidement évolué au cours des dernières années. Les logiciels antivirus ont dû adapter leurs techniques de détection, passant de la simple analyse de signatures à la combinaison de multiples méthodes de protection. Bien que les méthodes de tests aient évolué, elles présentent plusieurs limitations qui pourraient être adressées en partie par la méthode proposée dans le cadre du présent travail, soit le recours aux essais cliniques, tels qu'utilisés dans le domaine médical.

CHAPITRE 3

MÉTHODOLOGIE EXPÉRIMENTALE DE L'ESSAI CLINIQUE

Ce chapitre détaille la solution proposée ainsi que la méthodologie utilisée. Nous présentons dans un premier temps les notions d'éthique et de confidentialité relatives à l'expérience. S'en suit une description détaillée de l'équipement utilisé ainsi que du protocole expérimental. Finalement, nous détaillons les différents coûts reliés à la réalisation de l'expérience.

3.1 Description sommaire

La réalisation de l'expérience implique la participation de 50 volontaires pour une période de quatre mois. Ces derniers doivent, en début d'expérience, acheter à faible coût un ordinateur portable de qualité. Les ordinateurs utilisés dans le cadre de l'expérience sont préalablement configurés afin d'y installer un logiciel antivirus commercial ainsi que plusieurs scripts et applications. Ces derniers permettent de recueillir des informations sur la santé des ordinateurs ainsi que sur le comportement des utilisateurs. De plus, tous les ordinateurs portables sont connectés à un serveur central situé à Polytechnique qui permet le suivi en temps réel de leur état d'infection. Durant la période de l'expérience, les volontaires doivent se présenter à un total de quatre rencontres mensuelles pour lesquelles ils reçoivent une compensation financière pouvant atteindre la valeur d'achat de l'ordinateur. Lors de ces rencontres, les participants doivent remplir un questionnaire qui permet d'obtenir leur avis quant au logiciel antivirus ainsi que des informations sur leur utilisation de l'ordinateur. Les rencontres mensuelles permettent aussi de recueillir les différentes informations compilées par les scripts et de déterminer si l'ordinateur est infecté. En cas d'infection, des données supplémentaires sont collectées, sous consentement du participant, afin de déterminer la source et le moyen d'infection.

3.2 Certification du protocole

L'étude a été approuvée par le Comité d'évaluation des risques informatiques (CÉRI) ainsi que par le Comité d'éthique de la recherche (CÉR) de l'École Polytechnique de Montréal. Nous présentons dans la section qui suit les différents éléments qui ont été pris en considération lors du processus de certification.

3.2.1 Évaluation des risques informatiques du projet

L'expérience ne présente aucun risque additionnel à celui de l'usage normal d'un ordinateur portable pour les participants. Advenant que le logiciel antivirus détecte des infections, il s'occupera de les neutraliser. Toutefois, dans l'éventualité ou un de nos outils diagnostiques trouvent des infections sur l'ordinateur d'un participant, une procédure à suivre est donnée au participant afin qu'il puisse lui-même neutraliser la menace. De plus, afin de garantir un maximum de sécurité par l'antivirus, les serveurs du laboratoire utilisés pour distribuer les mises à jour de l'antivirus sont déployés dans un environnement de haute disponibilité afin de ne pas exposer les ordinateurs portables à un risque supplémentaire.

Étant donné que l'expérience requiert la manipulation de fichiers malveillants, des précautions particulières ont été prises afin de protéger les infrastructures de l'université. Tous fichiers identifiés comme étant potentiellement malveillants ont été chiffrés au préalable pour ensuite être entreposés dans la zone de haute sécurité du laboratoire.

3.2.2 Éthique et confidentialité

Toutes les données statistiques et brutes générées durant l'expérience sont anonymes étant donné qu'elles sont seulement identifiées par le numéro unique de l'ordinateur, dont seule la personne responsable du projet détient l'association avec l'identité du participant. Il n'est donc aucunement possible de pouvoir associer les informations récoltées avec l'identité d'un sujet.

La politique de sécurité du Labortoire SecSI, qui comprend toutes les mesures nécessaires en termes de sécurité physique, logique et du personnel, a été appliquée dans la protection des données récoltées. Toutes les informations recueillies au cours de l'expérience sont conservées dans un cabinet localisé dans la zone rouge du laboratoire, laquelle est sécurisée avec trois facteurs d'authentification (lecteurs biométriques, NIP, carte d'identification). Cette zone est complètement isolée d'Internet et du réseau de l'université. Seuls les membres du personnel impliqués dans le cadre du projet sont autorisés à avoir accès aux données. La politique du laboratoire est aussi appliquée pour la supression de toutes données reliées à l'expérience. Cette politique s'applique tant au format papier qu'au format électronique et est conforme aux standards du gouvernement canadien en matière de suppression de données.

L'utilisation des informations récoltées au cours de l'expérience se limite à l'atteinte des objectifs de recherche du projet. Cependant, dans les circonstances où les lois applicables l'obligent, comme par exemple la découverte fortuite d'informations pouvant mener une personne raisonnable à penser qu'un crime a été commis ou va être commis, nous aurions rapporté ces informations aux autorités compétentes.

3.3 Équipement

Nous avons remis le même ordinateur portable à tous les participants, soit un Toshiba Satellite C650D-06Q. Ces ordinateurs ont été préalablement configurés selon les mêmes spécifications. Le système d'exploitation Windows 7 Édition Familiale ainsi que le logiciel antivirus OfficeScan de Trend Micro ont été installés. Les outils diagnostiques suivants ont de plus été installés : Hijackthis, ProcessExplorer, Autoruns, SpyBHORemover, SpyDLLRemover, Tshark, WinPrefetchView, WhatChanged, Sigcheck ainsi que des scripts Perl que nous avons développés pour les besoins de l'expérience. Ces scripts sont entre autres destinés à exécuter automatiquement les différents outils diagnostiques en plus de collecter et compiler des informations statistiques sur la configuration et l'utilisation du système ainsi que l'environnement dans lequel il est utilisé.

Suite à la configuration d'un premier ordinateur portable, nous avons collecté différentes informations sur son état initial tel que le hachage de tous les fichiers présents ainsi que des informations concernant leur signature, la liste des processus actifs, la liste des programmes configurés pour démarrer à l'ouverture de l'ordinateur, la liste des différentes clés de registres, la liste des *Browser Helper Objects* (BHO), la liste des fichiers chargés lors du démarrage de l'ordinateur et la liste des fichiers prefetch. Ces données initiales ont été copiées sur l'ordinateur en question afin de servir de référence.

Une image de cet ordinateur portable a été crée, ce qui a permis de déployer cette dernière sur l'ensemble des ordinateurs. Nous avons attribué à chaque ordinateur un numéro d'identification unique variant entre 1 et 50, pour ensuite connecté manuellement l'ordinateur à un serveur central situé à Polytechnique afin d'assurer un suivi en temps réel de l'antivirus et de l'état d'infection de l'ordinateur.

3.4 Collecte des données

Nous avons receuilli trois types d'informations durant le projet, soit des données personnelles à propos des sujets (nom, numéro de téléphone, etc.), des données statistiques anonymes et des données brutes.

3.4.1 Données personnelles

Nous avons collecté en début d'expérience des informations personnelles concernant les participants afin d'assurer le bon déroulement de cette dernière. Ces informations sont limitées aux prénom et nom du participant, à son adresse courriel, son numéro d'ordinateur portable et son numéro de téléphone. Ces données permettent entre autres de contacter les participants durant l'expérience et de leur fournir des informations concernant leurs rencontres. Ces

informations sont conservées pour une période de 1 an après la fin de l'expérience après quoi elles seront détruites selon la Politique du laboratoire Secsi.

3.4.2 Données statistiques

Des jeux de données statistiques anonymes sont recueillis une fois par mois par des scripts développés pour les fins de l'expérience. Ces données incluent :

- la liste des applications installées ainsi que des applications pour lesquelles des mises à jour sont disponibles et ce, pour chaque jour ;
- le nombre de sites web visités par jour et ce, pour chaque navigateur web ;
- le nombre de sites web visités mensuellement par catégories de site pour chaque navigateur ;
- le nombre et le type de fichiers téléchargés par jour ;
- la liste mensuelle des plugins installés pour chaque navigateur ;
- le nombre d'heures par jour durant lesquelles l'ordinateur est connecté à Internet ;
- le nombre d'heures par jour durant lesquelles l'ordinateur est allumé ;
- le nombre de localisations différentes à partir desquelles l'ordinateur se connecte chaque jour ;
- la liste des différents serveurs DNS auxquels l'ordinateur se connecte chaque jour.

Ces différentes données servent à déterminer s'il existe des relations entre les variables étudiées, telles que le comportement de l'utilisateur, la configuration du système ou encore l'environnement, et le nombre d'infections détectées chez les participants.

3.4.3 Données brutes

Les données brutes sont seulement recueillies dans le cas où le logiciel antivirus trouve une infection, ou si nous soupçonnons qu'un ordinateur est infecté. Dans un tel cas, une autorisation spéciale est demandée au sujet afin de pouvoir récolter l'historique de navigation ainsi que les détails du trafic réseau du mois associé. Ces informations supplémentaires permettent de déterminer quelle est la source de l'infection et le moyen que cette dernière a utilisé pour infecter l'ordinateur portable. Tout comme pour les jeux de données statistiques, les données brutes sont anonymes puisqu'elles sont identifiées par le numéro du portable associé. Ces données peuvent toutefois comprendre des informations personnelles qui pourraient permettre à un tiers malveillant d'identifier le sujet. Afin d'éviter une telle situation, l'accès aux données brutes est hautement contrôlé. De plus, tout commes les données personnelles, les données brutes sont conservées pour une période maximale de 1 an suivant la fin de l'expérience.

3.5 Protocole expérimental

3.5.1 Recrutement

Nous avons en début d'expérience recruter 50 sujets avec comme seul critère d'inclusion le fait d'être majeur. Le recrutement des volontaires s'est principalement déroulé sur l'ensemble du campus de l'Université de Montréal, incluant l'École Polytechnique de Montréal et HEC Montréal.

Une campagne d'affichage (voir Annexe A) a eu lieu sur l'ensemble du campus universitaire et une annonce a été placée dans le journal étudiant de l'université. Bien que l'affichage ait principalement eu lieu sur le campus de l'Université de Montréal, la participation était ouverte à tous. Un site web (voir Annexe B) a été développé afin de présenter aux candidats intéressés le déroulement de l'expérience. Afin de faciliter le recrutement, un formulaire d'inscription a été mis en ligne. Les personnes intéressées n'avaient qu'à remplir quelques questions de base afin de soumettre leur candidature.

Les volontaires ont été classés en fonction de leur groupe d'âge, de leur genre, de leur statut (étudiant, employé, sans emploi) et de leur domaine d'activité. Sur les 131 demandes reçues, nous avons retenu 50 candidats en fonction de leur profil de sorte à obtenir un échantillon varié et représentatif de la population dans les limites des candidatures reçues.

Sur les 50 participants sélectionnés, 30 sont de genre masculin et 20 sont de genre féminin. Les participants ont été classés selon 7 groupes d'âge, soit 18-24, 25-30, 31-35, 36-40, 41-45, 46-50 et plus de 51 ans. Chaque groupe contient respectivement 19, 12, 7, 4, 1, 3 et 4 sujets. 32 participants sont des étudiants, 15 sont des travailleurs et 3 sont sans emploi. En ce qui concerne le domaine d'activité, 13 participants sont dans le domaine de l'informatique, 10 en sciences appliquées, 5 en sciences pures, 9 en science de la santé, 10 en sciences humaines, 1 en arts et lettres et 2 ont répondu qu'ils étaient dans un autre domaine.

3.5.2 Rencontres

L'expérience requiert des participants qu'ils se présentent à un total de cinq rencontres : une rencontre initiale ainsi que quatre rencontres mensuelles. La sous-section qui suit présente le déroulement général des rencontres. Pour plus de détails, le lecteur est invité à se référer au protocole de la rencontre initiale (voir Annexe I) et au protocole des rencontres mensuelles (voir Annexe J).

Les participants étaient invités à réserver une plage horaire via un calendrier de réservation hébergé sur nos serveurs. Le système de gestion *Meeting Room Booking System* a été utilisé afin de créer le calendrier en ligne.

Afin d'encourager les participants à assister à l'ensemble des rencontres, une indemnité

compensatoire est attribuée en fonction des rencontres auxquelles ils assistent. Une somme de 50$ est associée aux trois premières rencontres et une somme de 100$ est attribuée pour la dernière rencontre. Le participant qui assiste à la totalité des rencontres se voit remettre un bonus de 150$, pour un montant total de 400$, soit 50$ de plus que le prix d'achat de l'ordinateur portable.

Rencontre initiale

La première rencontre est de courte durée et sert principalement à présenter en détails le projet au participant et à lui remettre son ordinateur portable.

Chaque candidat retenu doit signer le formulaire de consentement (voir Annexe C) afin d'officialiser sa participation à l'étude et d'en accepter les termes et conditions. Par la suite, un ordinateur portable est vendu au participant au prix de 350$ et un reçu lui est émis. L'option de la vente a été retenue pour des raisons légales en plus de favoriser l'appropriation de l'ordinateur par le participant qui sera plus enclin à le personnaliser si ce dernier lui appartient. Cependant, les participants ne doivent en aucun cas reformater l'ordinateur, supprimer les logiciels et outils installés ainsi que les données collectées au cours de l'expérience et installer un autre logiciel antivirus.

Le participant est par la suite invité à remplir un questionnaire initial (voir Annexe D) destiné à reccueillir des informations générales sur son profil telles que son genre, son groupe d'âge, son statut, son domaine d'activité ainsi que son niveau d'expertise en informatique. Pendant ce temps, quelques dernières configurations sont apportées à l'ordinateur et ce dernier est connecté au serveur central de l'antivirus.

Rencontres mensuelles

Lors des rencontres mensuelles, les participants sont invités à remplir un questionnaire en ligne (voir Annexe E). Ce questionnaire vise à connaître l'avis du participant à l'égard du logiciel antivirus intallé ainsi qu'à obtenir un portrait global de son utilisation de l'ordinateur portable. Durant ce temps, les différentes données statistiques compilées au cours du dernier mois sont récoltées.

Au cours de chaque rencontre, l'ordinateur est analysé afin de détecter la présence de logiciels malveillants qui auraient échappés à l'antivirus. À cet effet, les outils suivants sont utilisés :

- Process Explorer : permet d'obtenir la liste complète des processus actifs sur l'ordinateur ;

- Autoruns : permet d'obtenir la liste complète des programmes exécutés automatiquement au démarrage de l'ordinateur ou à l'ouverture d'une session ;
- Hijackthis : permet d'obtenir entre autres la liste des programmes et des services exécutés au démarrage de l'ordinateur, la liste des BHO, les redirections dans le fichier hôte, etc. ;
- Sigcheck : analyse les images exécutables présentes sur l'ordinateur et vérifie si ces dernières sont signées ou non ;
- SpyBHORemover : dresse la liste des BHO installés et les classifie selon quatre catégories (dangereux, suspect, sécuritaire, non-classifié) ;
- SpyDLLRemover : dresse la liste des DLL installées et les classifie en trois catégories (dangereux, sécuritaire, non-classifié) ;
- Whatchanged : permet d'obtenir la liste complète des différentes clef de registres Windows ;
- Winprefetchview : permet d'obtenir la liste et le contenu des fichiers prefetch crées par Windows. Ces informations permettent de connaître quels sont les fichiers utilisés par les différentes applications exécutées, ainsi que la liste des fichiers chargés automatiquement au démarrage de Windows ;

Les différents fichiers analysés sur l'ordinateur sont par la suite classifiés en quatre catégories : dangereux, suspect, sécuritaire et non-classifié. Tous les fichiers marqués comme étant dangereux, suspects ou non-classifiés font l'objet d'une analyse supplémentaire afin de déterminer si ces derniers peuvent être considérés comme étant dangereux. Le service en ligne VirusTotal est utilisé afin d'analyser les fichiers par plusieurs moteurs antivirus ainsi que l'outil en ligne Anubis, qui permet d'analyser le comportement d'exécutables Windows.

Advenant le cas où un fichier malveillant ou potentiellement malveillant est identifié par notre protocole ou par l'antivirus, le participant doit remplir un questionnaire additionnel (voir Annexe F). Ce questionnaire vise à obtenir plus d'informations sur la source et le moyen par lesquels l'ordinateur aurait été infecté ainsi que sur les changements de comportement qui auraient pu être observés au niveau de l'ordinateur. Un formulaire de consentement supplémentaire (voir Annexe G) est présenté au participant afin d'obtenir l'autorisation de collecter l'historique de navigation et le journal des communications réseau associées au mois de la détection ainsi que le ou les fichiers malicieux. Sous l'accord du participant, le ou les fichiers malicieux sont préalablement chiffrés pour ensuite être copiés avec le reste des données et entreposés dans la zone de haute sécurité du laboratoire. Ces données brutes permettront par la suite de retracer la source et le moyen exact de l'infection.

Rencontre finale

La dernière rencontre est similaire aux rencontres mensuelles : les participants doivent remplir le questionnaire mensuel pendant que leur ordinateur est analysé et que les données sont collectées.

Toutefois, les participants doivent en plus remplir un questionnaire final (voir Annexe H) qui permet d'obtenir des informations sur leur expérience globale au cours de l'étude ainsi que sur les activités qui auraient pu influencer la nature de leurs résultats.

Un protocole est remis aux participants afin de leur expliquer comment arrêter la collecte automatique des données et supprimer les données enregistrées tout au long de l'expérience. Bien que les participants soient libres de supprimer les données collectées, ils sont invités à les conserver pour une période additionnelle de 3 mois. Ce délai nous permettrait de contacter les participants afin de collecter des informations additionnelles advenant le cas où une analyse supplémentaire serait requise.

En ce qui concerne les outils diagnostiques installés sur les ordinateurs portables dans le cadre de l'expérience, le choix est donné aux participants de les conserver ou de les supprimer une fois l'expérience complétée.

3.6 Coûts

Compte tenu que l'étude implique un achat majeur d'équipement et la participation de volontaires, les coûts du présent projet diffèrent de ceux associés à des expériences réalisées en laboratoire et méritent une discussion.

3.6.1 Dépenses initiales

Afin de fournir un ordinateur portable à tous les participants, 50 machines sont achetées au prix de 375$ chacune, pour un total de 18 750$. Cet achat est partiellement assumé par les participants qui doivent acheter leur ordinateur au prix de 350$, soit bien au deça du prix marchand estimé à 475-500. Le coût initial de l'expérience revient donc à 1 250$.

3.6.2 Dépenses de fonctionnement

L'embauche de personnel a été requise afin d'aider à la collecte des données lors des rencontres mensuelles. Un étudiant a été employé durant 95 heures pour un coût total de 1 800$. Un cellulaire a aussi été acheté pour les fins de l'expérience afin de pouvoir donner un numéro de téléphone dédié aux participants. L'achat ainsi que les frais d'utilisation s'élèvent à environ 200$. L'ensemble des dépenses de fonctionnement s'élève donc à 2 000$.

Cependant, ce montant ne prend pas en compte le temps investi dans la préparation et le suivi du projet par les différents chercheurs impliqués. En moyenne, 25 heures ont été dédiées aux rencontres initiales, 400 heures aux rencontres mensuelles et 25 heures ont été consacrées à du soutien technique envers les participants, ce qui donne un total de 450 heures.

3.6.3 Indemnités des participants

Chaque participant s'est vu offrir une license d'utilisation d'un an du logiciel antivirus utilisé, gracieuseté de la compagnie antivirus. Les utilisateurs ont aussi reçu des indemnités compensatoire de 50$ pour les 3 premières rencontres, 100$ pour la quatrième rencontre et un bonus de 150$ pour avoir assisté à l'ensemble des rencontres. Ainsi, les participants avaient la possibilité de recevoir jusqu'à 400$, ce qui donne un montant maximal de 20 000$. Sur l'ensemble des 50 participants, 47 ont assisté à la première rencontre, 49 à la deuxième, 49 à la troisième et tous se sont présentés à la dernière rencontre, totalisant un montant de 19 150$.

3.7 Conclusion

Nous avons dans ce chapitre exposé les notions de risques informatique et d'éthique qui ont été prises en compte afin d'assurer la confidentialité des participants et la protection des données collectées. Les différentes étapes du protocole expérimental ont été expliquées ainsi que le détail des coûts associés à la réalisation de l'étude. À cet effet, le coût total de l'expérience incluant les dépenses initiales, les dépenses de fonctionnement et les indemnités s'élève à environ 22 400$.

CHAPITRE 4

RÉSULTATS ET DISCUSSION

Ce chapitre présente les résultats obtenus suite à la réalisation de l'expérience décrite au chapitre 3. Nous présentons dans un premier temps les résultats permettant d'évaluer le logiciel antivirus utilisé. Nous analysons les menaces détectées par l'antivirus ainsi que les menaces manquées. Puis, nous présentons les résultats des évaluations des utilisateurs à l'égard du logiciel antivirus. Dans un second temps, nous analysons l'impact de différents facteurs humains sur la probabilité d'infection des utilisateurs. Plus précisement, nous étudions l'impact des facteurs démographiques et des caractéristiques ainsi que du comportement usager de sorte à déterminer quels sont les facteurs de risque d'infection.

4.1 Évaluation de l'antivirus

Afin de répondre au premier objectif du projet, soit d'évaluer un logiciel antivirus dans un contexte réel d'utilisation, nous analysons les détections de l'antivirus ainsi que les menaces non détectées au cours des quatre mois de l'expérience. De plus, les évaluations des utilisateurs sont comptabilisées afin de dresser un portrait global de l'appréciation de l'antivirus par les usagers.

4.1.1 Menaces détectées par l'antivirus

Durant les quatre mois de l'expérience, 380 fichiers ont été détectés par l'antivirus chez 19 participants, soit 38% de la polution totale. Toutefois, certains de ces fichiers ont été l'objet de détections multiples et ce, pour le même utilisateur. Sans ces répétitions, 95 détections ont été comptabilisées. La Figure 4.1 montre que la répartition de ces 95 détections pour chaque mois est similaire, ce qui signifie que les utilisateurs ne semblent pas adopter des comportements plus à risque lorsqu'ils commencent à utiliser un nouvel ordinateur.

Chacune de ces 95 détections est classifiée selon le type de menaces en fonction des informations fournies par l'antivirus. La Figure 4.2 permet de constater que la majorité des détections ont été causées par des chevaux de Troie, alors que la proportion de virus et de logiciels publicitaires est faible.

Ces résultats sont similaires à ceux raportés par d'autres compagnies antivirus. À titre d'exemple, le rapport du premier trimestre 2012 de Panda Security Labs (2012) indique que les chevaux de Troie comptent pour la majorité des détections avec une proportion de

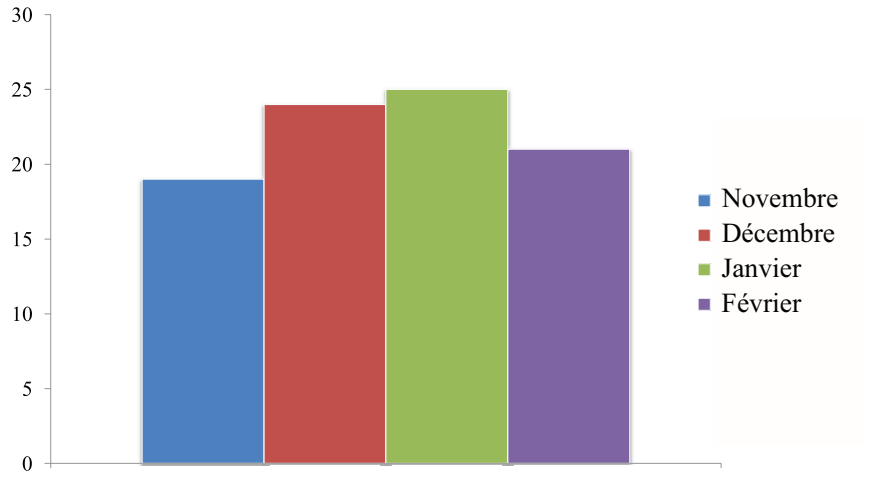


Figure 4.1 Nombre de détections par mois

63.30% alors que les vers, virus, logiciels publicitaires et les autres comptent respectivement pour 8.39%, 7.90%, 7.81% et 9.60%.

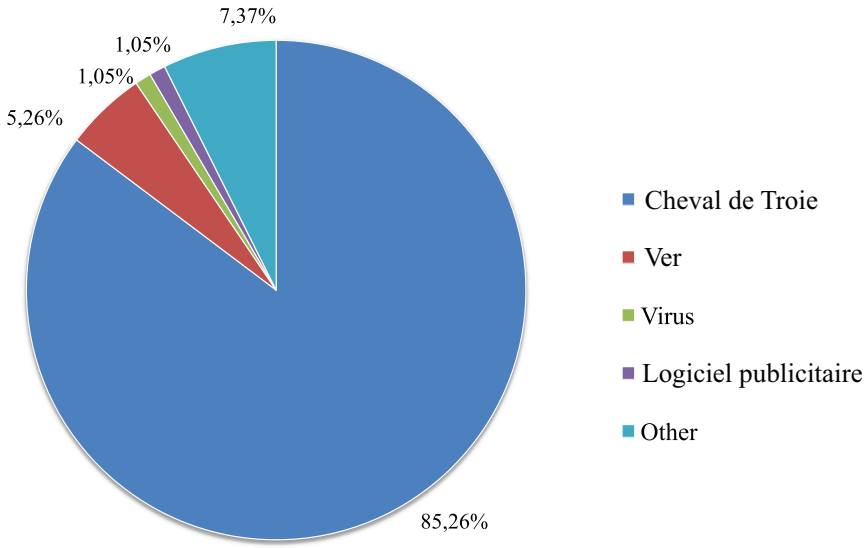


Figure 4.2 Classification des détections par type

La différence avec nos résultats peut en partie être attribuée à la méthode de classification utilisée. Par exemple, un fichier peut avoir été classé comme étant un cheval de troie par le logiciel antivirus utilisé alors qu'un autre produit antivirus l'aurait classé comme virus. De

plus, notre échantillon contient seulement 95 fichiers alors que les résultats de Panda Security sont basés sur un échantillon contenant des milliers de fichiers.

En ce qui concerne les mécanismes de détection, 93 des 95 fichiers ont été détectés par des analyses temps réel, et seulement deux fichiers ont été détectés par une analyse manuelle à la demande de l'utilisateur. Ces résultats sont considérés satisfaisants puisqu'un bon logiciel antivirus devrait détecter un maximum de menaces avec une intervention minimale de l'utilisateur. En réponse à ces 95 détections, l'antivirus a placé en quarantaine 78 fichiers, n'a pas réussi à placer en quarantaine 10 fichiers, à chiffrer deux fichiers et a identifié les 5 autres comme étant des menaces potentielles sans effectuer aucune action particulière.

Bien que l'analyse des sources et des moyens par lesquels ces fichiers se sont retrouvés sur les ordinateurs ne soit pas encore complétée, une analyse préliminaire basée sur le chemin vers les fichiers détectés nous a permis de conclure que 17 des 95 détections proviennent d'un dispositif de stockage externe tel qu'une clef USB, un DVD ou encore un disque dur externe.

4.1.2 Détections manquées

Le processus d'identification et de classification des fichiers potentiellement malveillants est basé sur le rapport des utilisateurs et sur l'analyse de journaux en provenance des différents outils de diagnostique utilisés.

Notre analyse a permis d'identifier 20 infections potentielles sur 12 ordinateurs, soit 24% des utilisateurs. L'outil Hijackthis a permis d'identifier 18 de ces détections, SpyBHORemover a permis d'en identifier une et la dernière infection a été reportée par l'utilisateur lui-même. Tous les fichiers identifiés comme étant potentiellement malveillants ont été collectés, sauf celui à l'origine de l'infection reportée par l'utilisateur. Bien que les outils diagnostiques utilisés aient été en mesure de repérer le fichier à la source de l'infection, ce dernier n'a pu être capturé puisqu'il se serait désinstallé automatiquement. Nous avons donc été en mesure de collecter 19 fichiers sur les 20 infections potentielles.

Les 19 fichiers collectés ont par la suite été analysés à nouveau par l'antivirus étudié afin de déterminer si ces derniers allaient être détectés à posteriori. Deux mois après la fin de l'expérience, aucun de ces fichiers n'a été identifié comme étant malveillant par l'antivirus. Chacun de ces 19 fichiers a de plus été analysé à posteriori par le service VirusTotal afin de comparer les résultats obtenus par différents produits antivirus et de comparer ces nouveaux résultats avec ceux obtenus précédemment au moment de leur détection. Cette analyse a permis de confirmer à nouveau la nature malveillante des fichiers. En complément, une recherche approfondie via Internet a été réalisée afin d'obtenir un maximum de détails sur les 20 infection potentielles. Suite à ces analyses, 2 des 20 menaces identifiées ont été classifiées comme étant sécuritaires, 7 comme des logiciels non désirés, 9 comme des logiciels publici-

taires, une comme un logiciel de sécurité falsifié et la dernière menace a été identifiée comme étant potentiellement malveillante.

Les neuf logiciels publicitaires détectés sont soit des BHO ou des barres d'outils. Dans tous les cas, ces logiciels n'ont pas été installés volontairement par les utilisateurs. Leurs effets sur l'ordinateur vont du changement de la page d'accueil des navigateurs web à la redirection des recherches ou à l'affichage de publicités. Une analyse supplémentaire sera nécessaire afin de déterminer si ces fichiers contiennent du code malveillant.

En ce qui concerne le logiciel de sécurité falsifié détecté, une brève analyse de l'ordinateur a permis de confirmer qu'il sagissait du faux antivirus Security Scanner. Selon le rapport de l'utilisateur, des fenêtres affichaient régulièrement que des logiciels malveillants étaient détectés sur l'ordinateur et toutes les applications lancées étaient automatiquement terminées, sauf les navigateurs web. Afin de nettoyer l'ordinateur, l'usager a été invité à payer par carte de crédit pour obtenir la version complète de l'antivirus, ce qui lui a fait soupçonner qu'il était possiblement infecté.

Au final, 18 menaces ont été détectées sur 10 utilisateurs, ce qui représente 20% des participants. En comparaison, l'analyse de 107 435 ordinateurs durant 55 jours a démontré que 32% des usagers sont infectés et ce, même s'ils ont un antivirus mis à jour (SurfRight (2009)). Bien que ces résultats ne peuvent être directement comparés à ceux de notre expérience, tous deux semblent démontrer que le risque d'infection est bel et bien réel et non négligeable même avec la présence d'un antivirus mis à jour régulièrement.

Bien que notre protocole ait permis d'identifier 18 menaces manquées par l'antivirus, nous avons l'intention de réaliser une analyse plus détaillée des différents journaux collectés, notamment la liste des registres Windows, afin d'identifier de nouvelles infections que nous n'aurions pas détectées.

4.1.3 Efficacité de l'antivirus

L'efficacité de l'antivirus (EA) évalué peut être obtenue en fonction du nombre de menaces détectées par le produit, soit les vrai-positif (VP), et du nombre de menaces qui n'ont pas été détectées, soit le nombre de faux-négatif (FN). Nous obtenons l'équation suivante :

$$EA = (\frac{VP}{VP + FN}) \tag{4.1}$$

Aux 95 détections uniques (VP), nous ajoutons les 18 menaces qui n'ont pas été détectées par l'antivirus (FN). Nous obtenons un total de 113 menaces auxquelles le produit antivirus aurait été exposé.

$$EA = (\frac{95}{95 + 18}) \tag{4.2}$$

$$EA = 0.8407 \tag{4.3}$$

L'antivirus étudié offrirait une efficacité de 84,07%. Plus présicement, ce résultat représente la sensibilité de l'antivirus à identifier correctement les logiciels malveillants. En comparaison, l'efficacité calculée est en deça des résultats obtenus par d'autres tests d'antivirus. À titre d'exemple, le dernier test réalisé par PC Security Labs (PC Security Labs (2013)) rapporte une efficacité de 99,99% pour Trend Micro et celui de AV-Comparatives (AV Comparatives (2013a)) rapporte une efficacité de 98,4%.

Alors que la majorité des produits antivirus performent avec des résultats similaires, variant généralement de 90% à 100%, nous croyons que le recours aux essais cliniques permettra d'obtenir des résultats plus réalistes en plus d'aider les compagnies d'antivirus à améliorer leurs produits.

4.1.4 Perception des utilisateurs

L'expérience réalisée vise à évaluer un logiciel antivirus au niveau de la détection mais aussi de l'appréciation globale des usagers. Différentes questions ont été adressées aux participants afin de connaître leur expérience et leur opinion quant au logiciel antivirus utilisé. Nous présentons dans la section qui suit une analyse préliminaire des résultats obtenus.

Niveau de protection procuré par l'antivirus

Les utilisateurs ont évalué le niveau de protection fourni par le logiciel antivirus installé sur leur ordinateur en attribuant une valeur allant de 1 à 10. Le Tableau 4.1 compare les différentes moyennes d'évaluation par mois entre les utilisateurs à risque, soit ceux qui ont déjà été exposés à une menace, et les utilisateurs à faible risque, soit ceux qui n'ont jamais été exposés à des menaces.

Tableau 4.1 Moyennes du niveau de protection

	À risque	Faible risque	Total
Novembre	8.4	7.7	7.8
Décembre	7.6	7.8	7.7
Janvier	6.9	7.6	7.5
Février	6.7	7.9	7.7
Total	7.4	7.7	7.7

En moyenne, sur les quatre mois de l'expérience, les participants ont évalué à 7,7/10 le niveau de protection fourni par l'antivirus. Plus précisement, les 23 utilisateurs ayant déjà eu au minimum une détection par l'antivirus ont évalué ce dernier en moyenne à 7,4/10 et ceux n'ayant jamais eu de détection ont évalué le niveau de protection à 7,7/10. Il ne semble donc pas y avoir de différence notable au niveau de la perception entre les utilisateurs ayant déjà eu des détections et ceux n'en ayant jamais eu.

Niveau d'interaction requis par l'antivirus

À la question visant à évaluer le niveau d'interaction requis par le logiciel antivirus, 36% des participants ont répondu qu'en moyenne, le niveau d'interaction requis est insuffisant, alors que 59% l'ont évalué adéquat et 3% ont jugé qu'il était trop élevé. Le Tableau 4.2 montre qu'il ne semble pas y avoir de différence notable entre les usagers à risque et ceux à faible risque quant à leur évaluation moyenne du niveau d'interaction.

Tableau 4.2 Niveau d'interaction moyen par groupe d'usagers

	À risque	Faible risque	Total
Insuffisant	37%	34%	36%
Adéquat	57%	61%	59%
Élevé	3%	3%	3%

La Figure 4.3 permet de suivre l'évolution des réponses au fil de l'expérience et de contaster qu'il n'y a pas de variation importante au niveau des résultats.

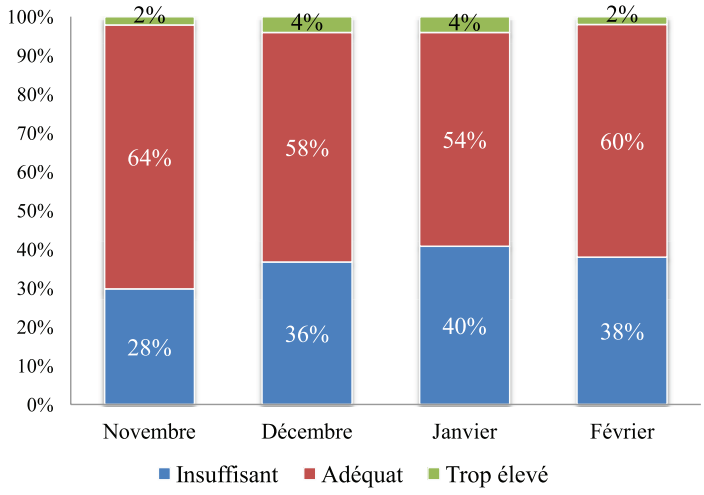


Figure 4.3 Niveau d'interaction par mois

Il convient de spécifier que le logiciel antivirus étudié s'adresse aux entreprises et qu'il est destiné à être installé sur des postes de travail. Par conséquent, l'antivirus est configuré afin de limiter les messages et donc l'interaction avec l'usager. Cette particularité de l'antivirus pourrait expliquer en partie pourquoi plus du tiers des utilisateurs ont jugé que le niveau d'interaction était insuffisant compte tenu du fait que les messages étaient limités à informer l'usager de la détection d'un fichier.

Sentiment à l'apparition d'une fenêtre de l'antivirus

Une des questions visait à déterminer le sentiment de l'utilisateur lors de l'apparition d'une fenêtre de l'antivirus. En moyenne, 56% des participants ont répondu se sentir réconforté à l'idée que l'antivirus fasse son travail, 18% se sont dit ennuyé d'être interrompu, 14% ont répondu être habitué et ne plus remarquer les fenêtres et 6% ont répondu qu'ils n'avaient jamais vu de fenêtre apparaître. La Figure 4.4 montre que la répartion des réponses est restée inchangée au cours de l'expérience.

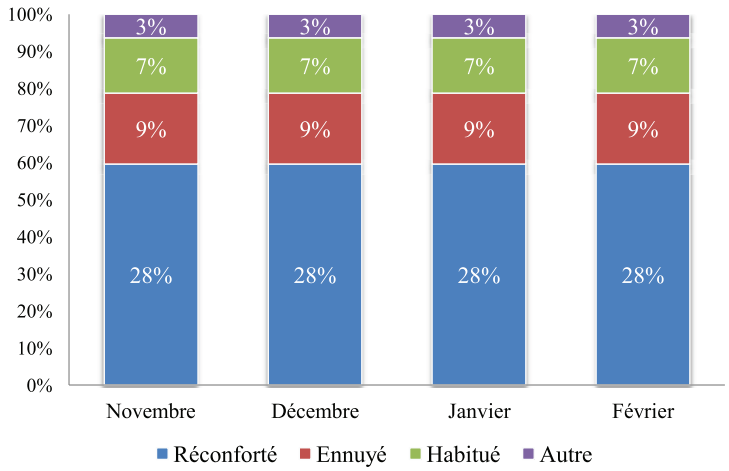


Figure 4.4 Sentiment à l'apparition d'une fenêtre de l'anti-virus

4.2 Impact des facteurs humains

Le second objectif consiste à déterminer quel est l'impact des facteurs humains sur le risque d'infection. Nous étudions l'impact des facteurs démographiques et des caractéristiques de l'utilisateur sur la probabilité d'infection ainsi que sur la fréquence d'infections. De plus, nous analysons le comportement de l'utilisateur afin de déterminer quels sont les comportements les plus à risque.

4.2.1 Facteurs démographiques et caractéristiques

Différentes informations concernant les participants sont collectées par l'intermédiaire du questionnaire initial lors de la première rencontre. Ces questions visent, entre autres, à déterminer le profil du participant à partir de son genre, son groupe d'âge, son statut social, son domaine d'activité et son expertise en informatique.

Ces données sont compilées afin de déterminer si certaines caractéristiques prédisposent les utilisateurs à un risque plus élevé d'infection. À cet effet, les participants sont divisés en deux groupes. Le premier contient les participants à risque, soit ceux qui ont eu au minimum une détection durant la durée totale de l'expérience, et le second contient les participants qui n'ont eu aucune détection au cours de l'expérience. Le Tableau 4.3 contient la proportion des utilisateurs pour l'ensemble de la population ainsi que pour le groupe à risque en fonction des différents facteurs démographiques et des caractéristiques.

Tableau 4.3 Proportion des utilisateurs pour chaque facteur

Facteur		Population totale	Population à risque	Population à faible risque
Genre	Homme	60%	61%	59%
	Femme	40%	39%	41%
Âge	18-24	38%	35%	40%
	25-35	38%	48%	30%
	36+	24%	17%	30%
Statut	Étudiant	64%	70%	59.3%
	Travailleur	30%	26%	33.3%
	Sans emploi	6%	4%	7.4%
Domaine d'activité	Informatique	26%	22%	30%
	Sciences naturelles	52%	48%	44%
	Arts et sciences	22%	30%	26%
Niveau technique en informatique	Élevé	18%	30%	7%
	Faible	82%	70%	93%
Expérience avec Windows 7	Oui	74%	78.3%	70.4%
	Non	20%	17.4%	22.2%
	Ne sais pas	6%	4.3%	7.4%
Expérience antérieure avec un antivirus	Oui	100%	100%	100%
	Non	0%	0%	0%

Le Tableau 4.3 laisse suggérer que l'âge et l'expertise en informatique pourraient potentiellement être des facteurs de risque. Afin de valider ces observations, une analyse statistique a été réalisée.

Nous avons dans un premier temps effectué une analyse des résidus, soit la différence entre

les valeurs observées et les valeurs estimées. Afin d'identifier la présence de données aberrantes à l'égard du nombre de détections uniques, nous avons analysé le graphique des résidus sur échelle de probabilité gaussienne. Ce graphique permet de vérifier si les données sont normalement distribuées, et par le fait même, d'identifier la présence de données abberantes.

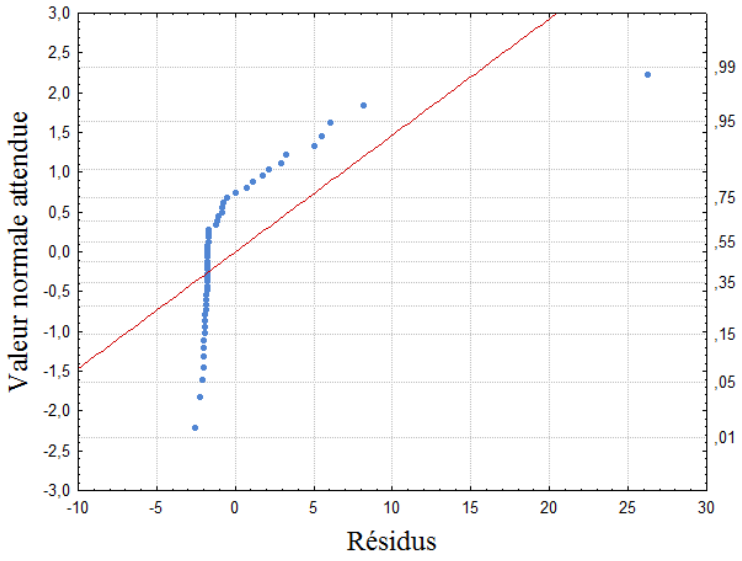


Figure 4.5 Résidus sur échelle de probabilité gaussienne

La Figure 4.5 permet de constater qu'un utilisateur se distingue des autres quant à son nombre de détections uniques. Une analyse détaillée des données a permis d'identifier l'utilisateur, soit l'usager 12, et de mieux comprendre les causes qui expliquent ces résultats. Le nombre élevé de détections chez cet utilisateur résulte en fait d'une infection qui n'a jamais été supprimée par l'antivirus. Les nombreux messages de l'antivirus ainsi que les problèmes causés par l'infection ont affecté l'usage de l'ordinateur par le participant. Qui plus est, l'usager a affirmé avoir prêté à plusieurs reprises l'ordinateur à des membres de sa famille. Par conséquent, cet usager peut être considéré comme étant marginal et ne sera pas pris en compte dans l'analyse.

La Figure 4.6 permet de constater que le nombre total de détections par utilisateur varie entre 0 et 12 détections si nous ne prenons pas en compte l'utilisateur 12. Les 49 participants considérés dans le cadre de l'analyse ont eu en moyenne 1,73 détections et ce nombre passe à 3,86 si nous considérons uniquement les usagers qui ont eu des détections.

Une régression logistique a été réalisée afin de déterminer si certaines caractéristiques ou facteurs démographiques influencent le risque d'infection, soit la probabilité d'avoir une détection ou plus. Ce modèle de régression binomiale est largement répandu dans plusieurs domaines afin de réaliser des analyses de risques. Appliqué à la médecine, il peut permettre

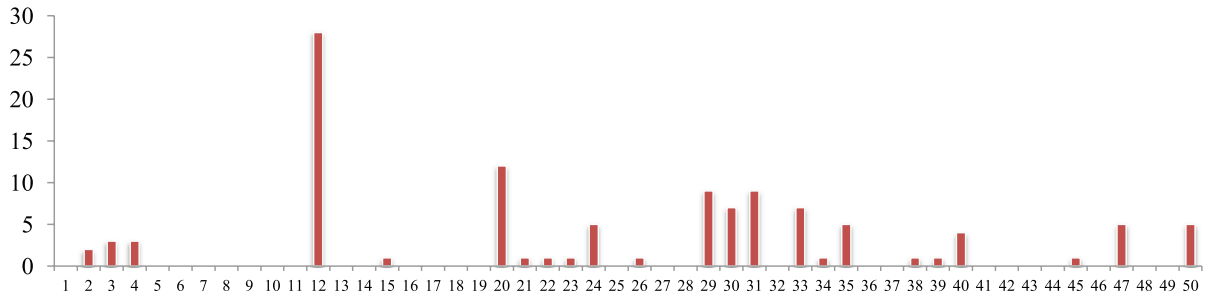


Figure 4.6 Nombre total de détections uniques par utilisateur

d'identifier les facteurs de risque d'une maladie. Cette méthode d'analyse est aussi très répandue dans le domaine des assurances et dans le domaine bancaire afin d'identifier les clients les plus à risque.

Dans le cadre de notre analyse, la variable dépendante, identifiée par 1 ou 0, représente le niveau de risque d'infection des utilisateurs. La valeur 1 est attribuée pour les utilisateurs à risque, soit ceux qui ont été exposés à une menace ou plus au cours de l'expérience, et 0 est attribué aux utilisateurs à faible risque, soit ceux qui n'ont pas été exposés à des menaces. Nous avons retenu le coefficient chi-carré (χ^2) qui permet de mesurer l'adéquation du modèle. Une valeur élevée indique que le modèle obtenu offre une bonne prédiction des données, c'est-à-dire que la déviance entre les valeurs prédites et observées est faible. La valeur p représente le niveau de signification, soit la probabilité que le résultat soit obtenu par chance. Les résultats sont considérés statistiquement significatifs si cette probabilité est inférieure à 5% ($p< 0.05$).

Tableau 4.4 Régression logistique comparant le risque d'infection

Facteur	χ^2	p
Genre	0.10	0.75
Âge	3.09	0.21
Statut	0.31	0.86
Domaine	0.47	0.79
Niveau d'expertise en informatique	**4.96**	**0.03**

Les résultats de la régression logistique présentés dans le Tableau 4.4 permettent d'identifier qu'un seul facteur, le niveau d'expertise en informatique, semble affecter la probabilité d'infection des usagers.

Nous avons de plus réalisé une analyse de la variance (ANOVA) à un facteur afin d'évaluer si le nombre de détections uniques est affecté par les caractéristiques et les facteurs démo-

graphiques des utilisateurs. Entre d'autres mots, nous avons voulu déterminer quels facteurs ont une influence sur la distribution de la variable dépendante, soit le nombre de détections uniques par usager. Les résultats de cette analyse sont présentés dans le Tableau 4.5.

Tableau 4.5 ANOVA à un facteur

Facteur	SS	MS	F	p
Genre	10.36	10.36	1.69	0.20
Âge	19.25	9.62	1.58	0.22
Statut	5.10	2.55	0.40	0.67
Domaine	21.56	10.78	1.79	0.18
Niveau d'expertise en informatique	4.41	4.41	0.70	0.41

La somme des carrés, *sum of squares* (SS) en anglais, représente la somme des carrés des écarts par rapport à la moyenne, alors que la moyenne quadratique, *mean squares* (MS) en anglais, est un estimé de la variance entre les groupes. Des résultats élevés pour ces valeurs est un signe qu'il existe une différence importante entre les groupes comparés. La valeur F permet de déterminer si la variation entre deux groupes est significative. Généralement, une valeur élevée ($F \gg 1$) laisse suggérer qu'il existe une différence importante entre la moyenne des groupes. Finalement, la valeur p permet de déterminer quelle est la probabilité que le même résultat soit obtenu par chance. Encore une fois, un résultat est considéré statistiquement significatif si p est inférieur à 5%. À cet effet, il semblerait qu'aucun des facteurs étudiés n'ait un impact significatif sur le nombre d'infection des utilisateurs.

Genre

La population totale inclut 30 hommes et 20 femmes, ce qui donne une proportion de 60% et 40% respectivement. Le Tableau 4.3 montre que la distribution des genres parmi les 23 utilisateurs classés comme étant à risque est de 61% pour les hommes et 39% pour les femmes. La régression logistique réalisée (Tableau 4.4) révèle qu'il n'y a pas de différence significative entre les hommes et les femmes quant au risque d'infection.

Nous avons examiné le nombre de détections uniques afin de déterminer s'il existe une différence entre les deux genres. La Figure 4.7 suggère qu'il y a une légère différence entre le nombre moyen de détections uniques. Toutefois, l'analyse de la variance (Tableau 4.5) révèle qu'il n'y a pas de différence significative au niveau des détections uniques entre les hommes et les femmes.

Plusieurs études ont adressé la question de la différence des genres en sécurité informatique, mais très peu ont étudié cette différence à l'égard du risque d'infection. Suite à un

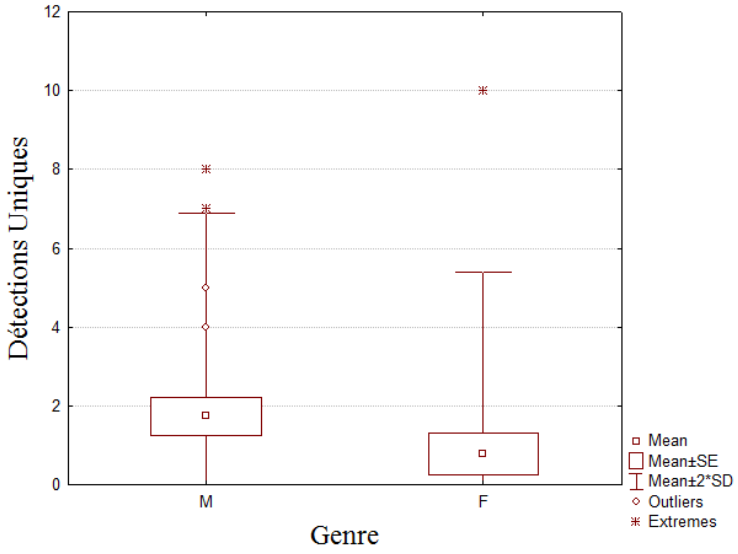


Figure 4.7 Boîtes à moustaches des détections uniques par genre

sondage réalisé auprès de 295 étudiants, Ngo et Paternoster (2011) ont trouvé que le genre n'était pas un facteur de risque significatif quant au risque d'infection par virus. Toutefois, ces résultats sont basés sur l'auto-évaluation des répondants et non sur des données réelles. Au meilleur de notre connaissance, aucune expérience empirique n'a été réalisée afin d'évaluer la différence entre les hommes et les femmes à l'égard de leur risque d'infection. À cet effet, notre analyse suggère qu'il n'y a pas de différence entre les deux genres.

Âge

Les utilisateurs ont été divisés en trois groupes selon leur âge. Le Tableau 4.3 montre que la proportion des 18-24 ans qui sont à risque est similaire à la proportion pour la population totale. Pour les 25-35 ans, la proportion des utilisateurs à risque (48%) est légèrement supérieure à celle de la population totale (38%). Et pour le groupe des 36 ans et plus, nous observons une faible diminution de 7% entre la proportion totale de la population (24%) et celle des utilisateurs à risque (17%). Les résultats de la régression logistique dans le Tableau 4.4 montrent qu'il n'y a pas de différence significative entre les différents groupes d'âge pour ce qui est du niveau de risque d'infection.

Nous avons par la suite étudié si le nombre de détections uniques était influencé par l'âge des participants. Bien que la Figure 4.8 laisse suggérer la présence d'une différence entre les groupes d'âge, les résultats obtenus ne sont pas significatifs.

D'autres études ont obtenus des résultats contraires, comme quoi l'âge des utilisateurs serait un facteur de risque. Sheng *et al.* (2010) ont réalisé une enquête en ligne auprès de 1001

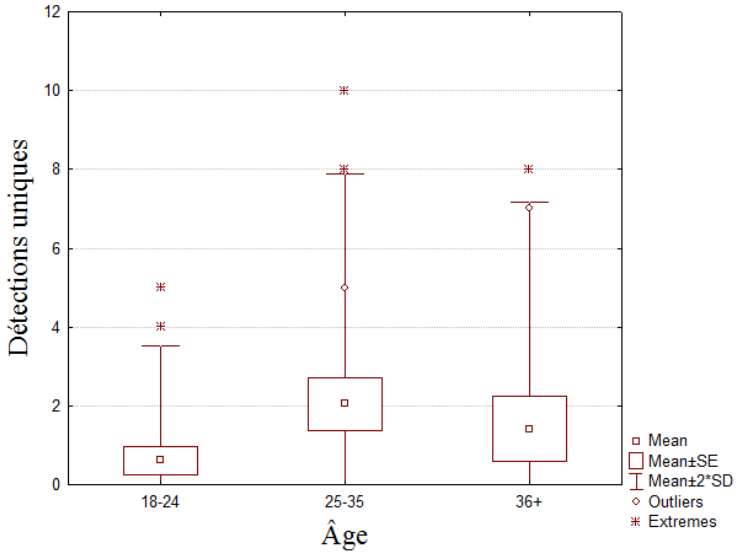


Figure 4.8 Boîtes à moustaches des détections uniques par groupe d'âge

usagers et ont trouvé que les 18-25 ans sont plus susceptibles aux attaques d'hameçonnage puisqu'ils ont, entre autres, un niveau d'éducation mois élevé, moins d'années d'expérience avec Internet et une plus faible aversion aux risques financiers ; Milne *et al.* (2009) ont conclu suite à un sondage mené auprès de 449 acheteurs en ligne que le comportement des usagers quant aux menaces informatiques varient en fonction de l'âge ; et Ngo et Paternoster (2011) ont mené un sondage basé sur l'auto-évaluation de 295 étudiants, ce qui leur a permis d'identifier que l'âge est un facteur de risque d'infection significatif et que les personnes âgées sont moins susceptibles d'être infectées par des virus informatiques.

Statut social

Les participants ont été classés en trois catégories distinctes en fonction de leur statut : étudiant, travailleur et sans emploi. Le Tableau 4.3 indique que la proportion entre la population à risque et la population totale est similaire pour chaque catégorie. La régression logistique (Tableau 4.4) vient confirmer ces résultats comme quoi le risque d'infection n'est pas plus important chez un groupe en particulier.

En ce qui concerne le nombre de détections uniques, la Figure 4.9 vient illustrer les variations selon le statut des participants. Les résultats de l'analyse de la variance (Tableau 4.5) montre qu'il n'existe pas de différence significative entre les différents statuts et donc, que le statut d'un utilisateur ne serait pas un facteur de risque.

Ces résultats viennent confirmer ceux obtenus par Ngo et Paternoster (2011) qui ont trouvé que le risque d'infection par virus informatiques est le même chez les étudiants, in-

dépendamment qu'ils aient un emploi à temps plein, à temps partiel ou qu'ils soient sans emploi.

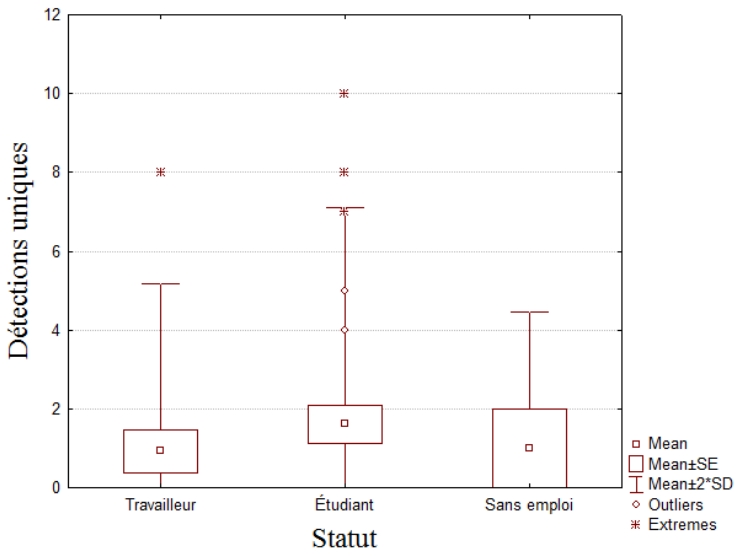


Figure 4.9 Boîtes à moustaches des détections uniques par statut

Domaine d'activité

Nous avons recruté les participants en fonction de leur domaine d'activité afin d'obtenir un échantillon représentatif de la population. Le Tableau 4.3 montre que 26% des participants sont en informatique, 52% en sciences naturelles et 22% en arts et sciences. Bien que les données du Tableau 4.3 suggèrent que les utilisateurs en arts et sciences semblent légèrement plus à risque, la régression logistique réalisée (Tableau 4.4) démontre qu'il n'existe pas de différence significative entre les trois groupes.

La Figure 4.10 présente le nombre moyen de détections uniques selon le domaine d'application. Fait intéressant, les participants en arts et sciences semblent avoir en moyenne un nombre de détections moins élevé que les autres groupes. Toutefois, l'analyse de la variance (Tableau 4.5) confirme que ces résultats ne sont pas statistiquement significatifs.

Solic et Ilakovac (2009) ont obtenus des résultats similaires. Ils ont interrogé un groupe de 19 médecins et 20 ingénieurs électrique afin de déterminer s'il existe une différence entre leur comportement usager et leur niveau de connaissances à l'égard de la sécurité informatique. À cet effet, l'analyse a permis de conclure que l'environnement de travail et l'expérience n'ont pas d'impact significatif sur le comportement des usagers hautement éduqués, ce qui vient supporter nos résultats.

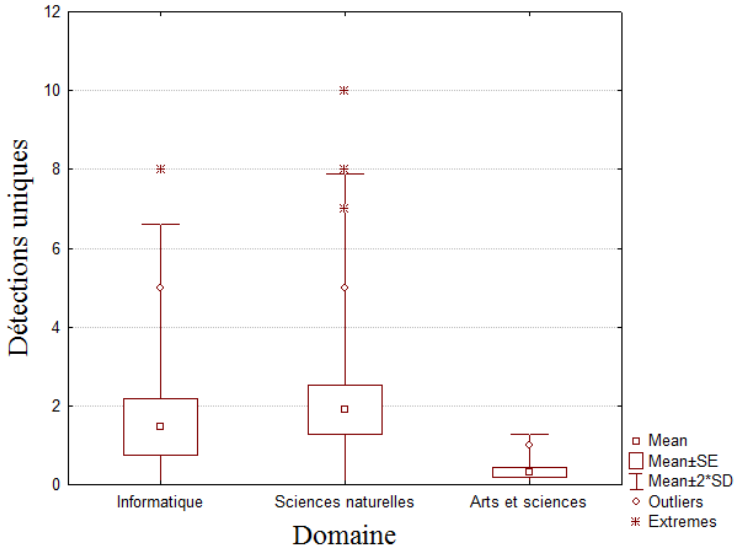


Figure 4.10 Boîtes à moustaches des détections uniques par domaine d'application

Expérience en informatique

Nous avons calculé le niveau d'expertise en informatique des participants en fonction de certaines tâches techniques déjà réalisées. Les utilisateurs ayant déjà crée un site Web, installé ou réinstallé un système d'exploitation ou ayant déjà configuré un réseau sans-fils ont été considérés comme ayant un niveau élevé d'expertise en informatique. Au total, 18% des participants ont été classés comme ayant un niveau élevé d'expertise.

En ce qui concerne l'utilisation de Windows 7, la majorité (74%) des participants ont répondu avoir déjà utilisé ce système d'exploitation, 20% ne l'ont jamais utilisé et 6% n'étaient pas en mesure de répondre. Au niveau de l'expérience antérieure avec un ou des antivirus, l'ensemble des participants ont répondu avoir déjà utilisé un antivirus.

Tel que présenté dans le Tableau 4.3, la proportion des utilisateurs ayant un niveau élevé d'expertise semble presque doubler entre la population totale et la population à risque. Ces résultats suggèrent que les utilisateurs ayant un niveau élevé d'expertise en informatique seraient potentiellement plus à risque d'être infectés. Notre régression logistique vient confirmer ces résultats comme quoi les utilisateurs ayant un niveau élevé d'expertise en informatique sont plus susceptibles d'être infectés. Nous avons calculé à partir des proportions données dans le Tableau 4.3 le rapport des chances, qui correspond au rapport entre la probabilité d'occurence d'une infection dans le groupe des usagers ayant un niveau d'expertise élevé et le groupe des usagers ayant un niveau d'expertise faible. Ce rapport indique que les utilisateurs ayant un niveau d'expertise élevé en informatique auraient 5,47 fois plus de chances d'être exposés à des menaces.

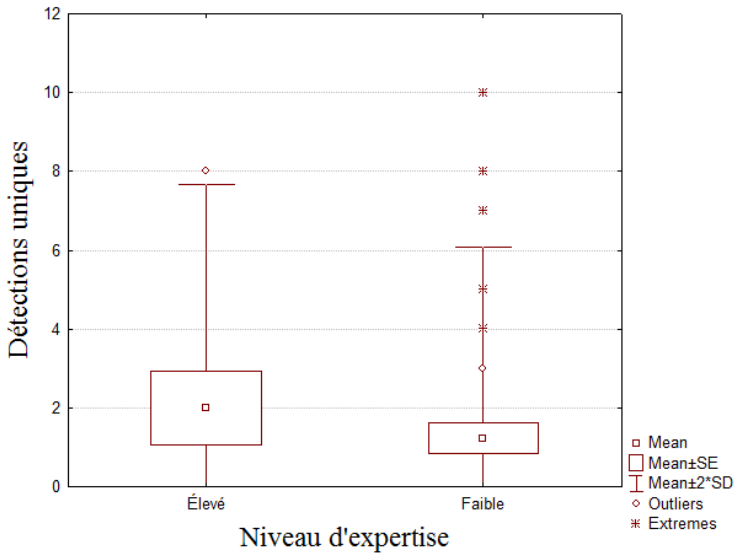


Figure 4.11 Boîtes à moustaches des détections uniques par niveau d'expertise

Comme illustré à la Figure 4.11, nous avons par la suite comparé le nombre de détections uniques entre les utilisateurs ayant un niveau élevé et ceux ayant un faible niveau d'expertise. Bien qu'une plus grande variance soit apparente pour les utilisateurs ayant un niveau d'expertise élevé, la différence ne semble pas significative (Tableau 4.5). Ainsi, les utilisateurs ayant un niveau élevé d'expertise en informatique semblent plus susceptibles d'être infectés, mais ils ne semblent pas pour autant avoir été exposés à un nombre plus élevé de menaces.

4.2.2 Comportement usager

Afin d'évaluer si certains comportements utilisateur entraînent une augmentation du risque d'infection, nous avons étudié l'impact des facteurs suivants : le navigateur Web le plus utilisé, le nombre d'applications installées, la durée totale d'activité système, le temps total de connexion, le nombre total d'hôtes contactés, le nombre total de sites Web visités et les catégories de sites visités.

Une approche similaire à celle décrite à la Section 4.2.1 a été utilisée. Nous avons premièrement effectué une régression logistique où les participants sont divisés en deux groupes : les utilisateurs à risque et les utilisateur à faible risque. Nous avons par la suite réalisé une régression générale afin d'évaluer s'il existe une relation entre les facteurs comportementaux analysés et le nombre de détections uniques. Une analyse de régression générale a été retenue au lieu d'une analyse de variance étant donné la nature continue des variables indépendantes.

Le Tableau 4.6 résume les différentes analyses où les résultats en gras sont considérés comme étant statistiquement significatifs ($p < 0.05$). Nous avons retenu le coefficient χ^2 pour

la régression logistique et le coefficient t, qui permet de déterminer le pouvoir explicatif d'une variable, pour la régression générale.

Tableau 4.6 Analyse statistique du comportement usager

	Régression logistique		Régression générale	
Facteur	χ^2	p	t	p
Navigateur Web	5.873	0.050	-	0.50
Nombre total d'applications installées	1.582	0.208	**2.289**	**0.027**
Durée totale d'activité système	1.093	0.296	0.630	0.531
Temps total de connexion	**7.325**	**0.007**	1.516	0.136
Nombre total d'hôtes contactés	**9.442**	**0.002**	**4.461**	**0.00003**
Nombre total de sites Web visités	**6.247**	**0.012**	1.713	0.093
Nombre de sites Web visités par catégorie				
Flux vidéo/MP3	**11.999**	**0.001**	1.919	0.061
Pair à pair (*peer-to-peer*)	**6.864**	**0.009**	4.918	0.00005
Infrastructure Internet	**7.469**	**0.006**	**4.000**	**0.0002**
Téléchargement logiciels	**14.326**	**0.000**	**5.293**	**0.000003**
Sports	**4.194**	**0.041**	1.380	0.395
Réseaux sociaux	**6.260**	**0.012**	0.860	0.395
Ordinateurs/Internet	**7.357**	**0.007**	**2.478**	**0.017**
Pari (*gambling*)	**4.998**	**0.025**	1.840	0.076
Pornographie	2.930	0.087	**2.440**	**0.020**
Illégal/Questionnable	0.022	0.881	1.863	0.095
Traducteur/page cache	1.689	0.194	**2.355**	**0.025**

Navigateur Web

Chaque mois, les participants devaient identifier le navigateur Web qu'ils ont le plus utilisé. Nous avons compilé ces résultats afin d'obtenir pour chaque utilisateur le navigateur Web qui a été le plus utilisé au cours de l'expérience. Le Tableau 4.7 résume l'usage des différents navigateurs au sein de la population totale et de la population à risque. Bien qu'une légère augmentation soit visible au niveau du navigateur Chrome, notre régression logistique révèle que le choix d'un navigateur en particulier n'affecte pas la probabilité d'infection des utilisateurs.

Nous avons aussi étudié si l'utilisation d'un navigateur en particulier influence le nombre de détections des utilisateurs. Les résultats de notre régression générale présentés dans le Tableau 4.6 viennent confirmer que l'utilisation d'un navigateur Web en particulier n'influence pas le nombre de détections. Par conséquent, le recours à un navigateurs Web à license propriétaire ou à code source ouvert n'exposerait pas l'usager à un risque addtionnel.

Tableau 4.7 Navigateur Web utilisé

Navigateur web	Population totale	Population à risque
Internet Explorer	30%	17.4%
Firefox	30%	26.1%
Chrome	40%	56.5%

Nombre d'applications installées

Nous avons collecté tout au long de l'expérience le nombre d'applications installées par participant tel qu'illustré à la Figure 4.12. Au total, 3 389 applications ont installées en quatre mois par les 50 usagers. Les participants ont installé entre 2 et 177 applications chacun, pour une moyenne de 68 applications.

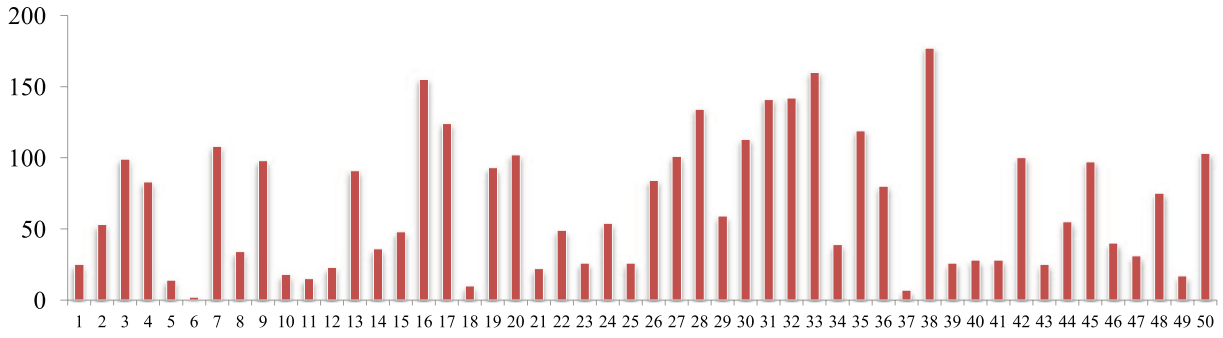


Figure 4.12 Nombre total d'applications intallées par utilisateur

Afin d'évaluer l'impact potentiel sur le risque d'infection, nous avons effectué une régression logistique, tel qu'indiqué dans le Tableau 4.6. Nos résultats suggèrent qu'installer un nombre élevé d'applications n'augmente pas la probabilité d'infection d'un usager.

La Figure 4.13 vient illustrer nos résultats, comme quoi les utilisateurs à risque n'ont pas installé plus d'applications que les utilisateurs à faible risque.

En complément, notre régression générale permet de voir que le nombre d'applications installées influence significativement le nombre de détections (Tableau 4.6). Ainsi, plus un utilisateur installe des applications, plus il est susceptible d'être infecté.

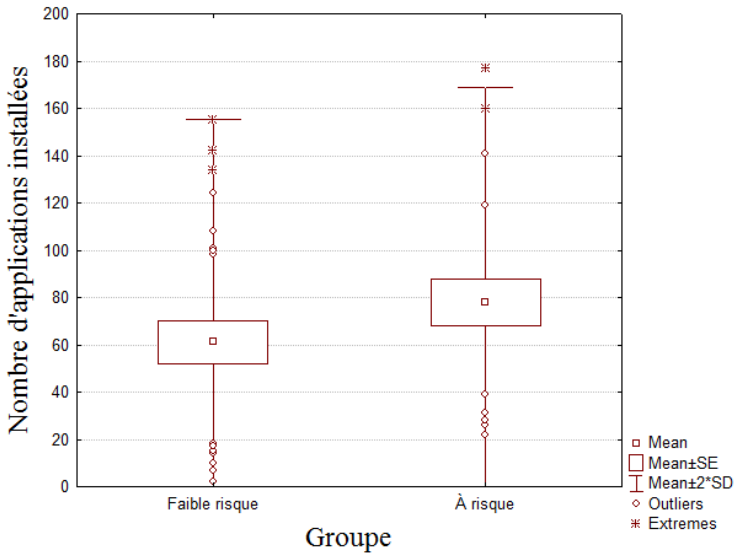


Figure 4.13 Boîtes à moustaches du nombre d'applications installées par groupe

Durée d'activité système

Le temps d'activité système a aussi été collecté pour chaque jour de l'expérience, ce qui nous permet d'obtenir la durée totale en heures pour les quatre mois de l'étude. La Figure 4.14 permet de constater que les ordinateurs sont restés allumés entre 109 et 2 882 heures pour la durée totale de l'expérience. Les participants ont laissé en moyenne 1 647 heures lors ordinateur allumé pour un total de 82 8332 heures pour l'ensemble des usagers.

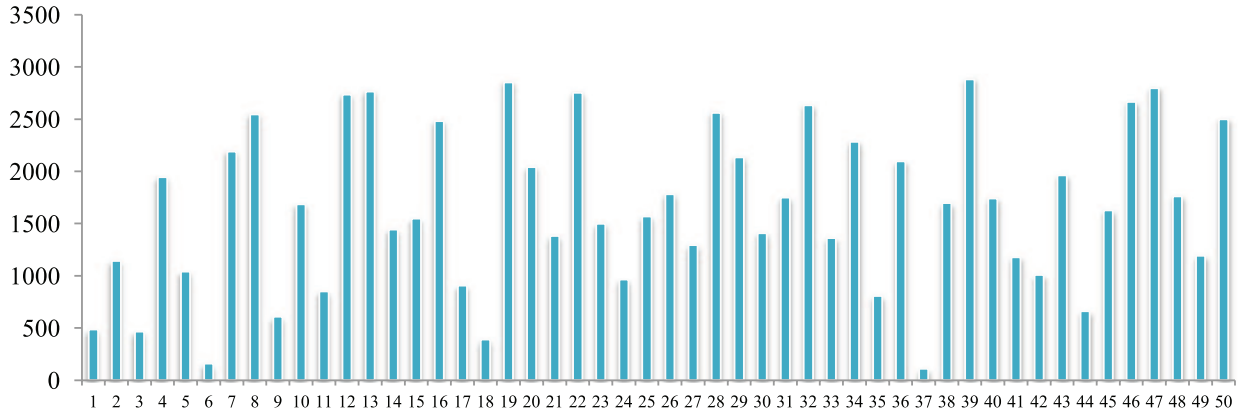


Figure 4.14 Temps total d'activité système en heures par utilisateur

Afin de déterminer si ce facteur influence le risque d'infection d'un utilisateur, nous avons

réalisé une régression logistique. Les résultats présentés au Tableau 4.6 indiquent que la durée d'activité système ne serait pas un facteur de risque quant à la probabilité d'infection.

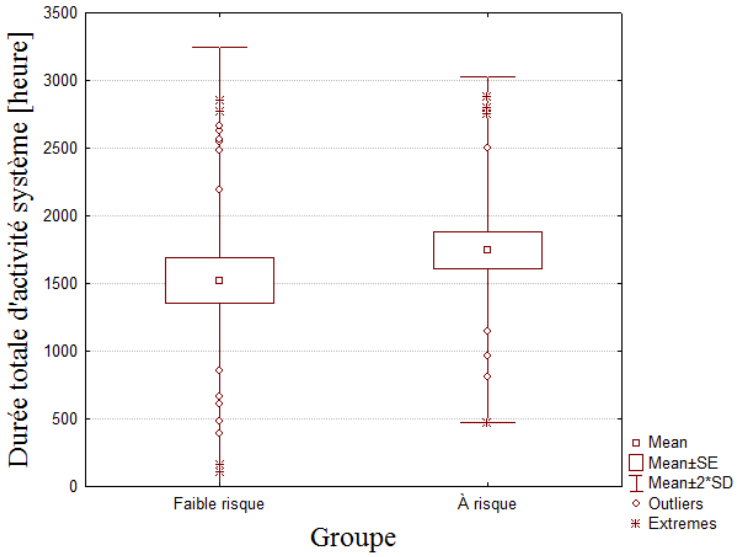


Figure 4.15 Boîtes à moustaches de la durée totale d'activité système

La Figure 4.15 vient illustrer que la durée moyenne d'activité du système entre les utilisateurs à risque et les utilisateurs à faible risque est similaire. Ces résultats sont confirmés par notre régression générale, comme quoi la durée d'activité système n'influence pas le nombre de détections. Ainsi, le fait d'avoir son ordinateur allumé durant plusieurs heures n'exposerait pas l'utilisateur à un risque d'infection plus élevé.

Temps total de connexion

Pour chaque jour, nous avons calculé le nombre d'heures durant lesquelles l'ordinateur portable a été connecté à Internet, ce qui nous a permis d'obtenir le temps total de connexion en heures. La Figure 4.16 montre que les usagers ont passé entre 12 et 992 heures en ligne pour la durée total de l'expérience. Au total, les participants ont passé 12 103 heures en ligne pour une moyenne de 242 heures.

Afin d'analyser l'impact sur le risque d'infection, nous avons effectué une régression logistique (Tableau 4.6) qui confirme que plus un utilisateur passe du temps en ligne, plus il augmente son risque d'être infecté. La Figure 4.17 vient illustrer la différence entre les utilisateurs à risque et les utilisateurs à faible risque quant à leur temps total de connexion.

Pour ce qui est de l'impact sur le nombre de détections, nos résultats obtenus par régression générale montrent que le temps de connexion n'influence pas directement le nombre de détections d'un utilisateur. Par conséquent, un utilisateur dont l'ordinateur est souvent

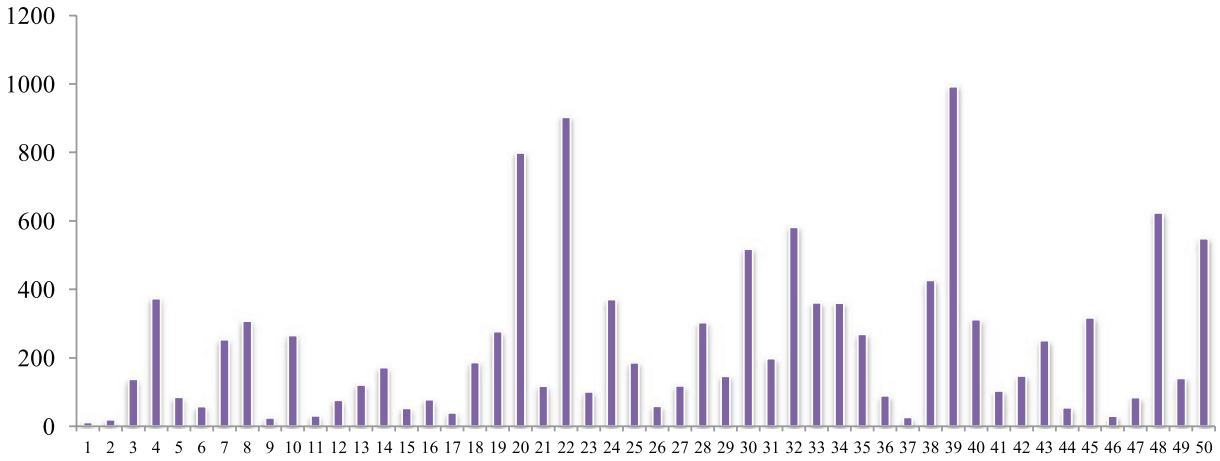


Figure 4.16 Temps total de connexion en heures par utilisateur

connecté à Internet augmente son risque d'être infecté. Cependant, il n'existerait aucune relation entre le nombre d'heures de connexion et le nombre de détections.

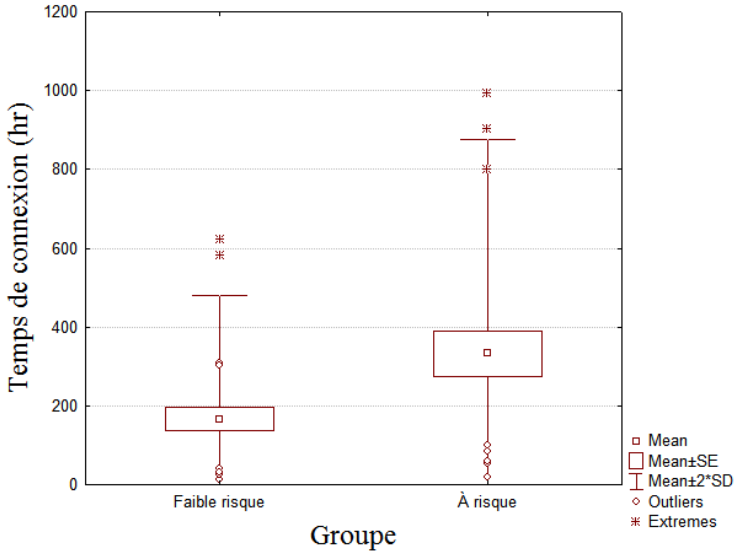


Figure 4.17 Boîtes à moustaches du temps total de connexion

Nombre total d'hôtes contactés

Nous avons été en mesure de calculer le nombre d'hôtes différents contactés par jour pour chaque utilisateur à partir des journaux générés par Tshark, une application qui permet de faire de la capture réseau. Ces données nous ont permis d'obtenir le nombre total d'hôtes

contactés par mois. La Figure 4.18 montre le nombre total d'hôtes contactés pour chaque utilisateur. L'axe des ordonnées a toutefois été limité à 300 000 afin d'offrir une meilleure vue d'ensemble. Le nombre d'hôtes différents contactés varie entre 18 et 1 508 833 hôtes par utilisateur pour une moyenne de 60 433 hôtes. Au total, 3 021 650 hôtes différents ont été contactés par les utilisateurs en quatre mois.

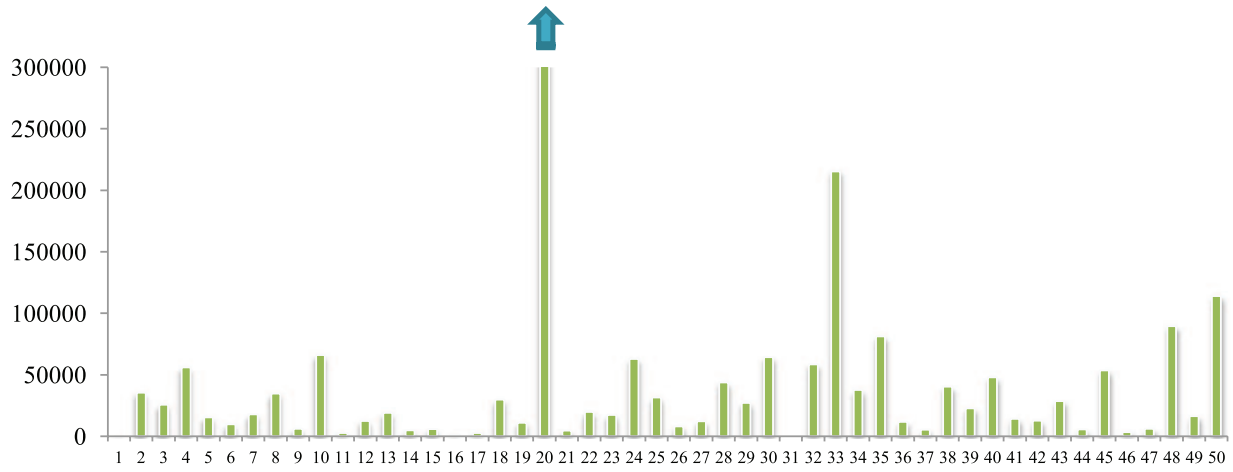


Figure 4.18 Nombre total d'hôtes contactés par utilisateur

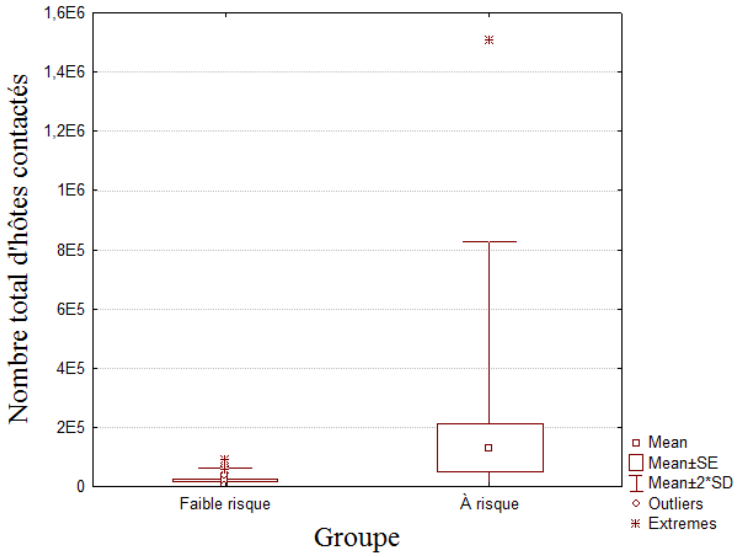


Figure 4.19 Boîtes à moustaches du nombre total d'hôtes contactés

Afin de déterminer si cette variable est un facteur de risque, nous avons réalisé une régression logistique (Tableau 4.6) qui indique que le nombre total d'hôtes contactés serait

un facteur de risque ($p < 0.05$). Ce résultat signifie que plus un utilisateur contacte un nombre élevé d'hôtes, plus il augmente ses risque d'infection.

Nous avons de plus réalisé une régression générale afin de vérifier si le nombre d'hôtes contactés influence le nombre de détections. À cet effet, notre analyse vient confirmer que cette variable serait un facteur de risque tel qu'illustré à la Figure 4.19.

Nombre de sites Web visités

Le nombre de sites Web visités a aussi été collecté afin d'étudier l'impact sur le risque d'infection. La Figure 4.20 présente le nombre total de sites Web visités pour chaque usager. Tel qu'illustré, les usagers ont visité entre 481 et 69 247 sites Web. Au total, ils ont visité 926 556 sites Web pour une moyenne de 18 531 sites Web.

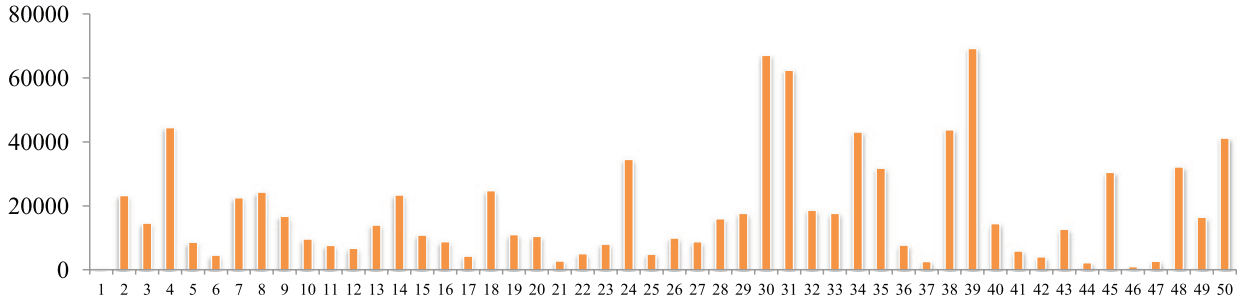


Figure 4.20 Nombre total de sites Web visités par utilisateur

La Figure 4.21 montre que les utilisateurs à risque ont visité beaucoup plus de sites Web que les utilisateurs dits à faible risque. La régression logistique (Tableau 4.6) confirme cette observation comme quoi le nombre de sites Web visités influence significativement le risque d'infection.

La régression générale ne confirme pas que le nombre de sites Web visités influence le nombre de détections chez l'usager. Toutefois, la p-value associée (0,09) est inférieure à 0,1, ce qui laisse suggérer que le nombre de sites Web visités pourrait être un facteur de risque à valider dans une étude de plus grande échelle.

Très peu d'expériences ont été réalisées à partir de données réelles afin de déterminer si le nombre de sites Web visités influence le risque d'infection. Carlinet *et al.* (2008) a étudié le trafic réseau généré par des centaines d'utilisateurs et a trouvé que ceux visitant un nombre élevé de sites Web s'exposaient davantage à générer du trafic malicieux et donc, à être infectés. Nos résultats viennent confirmer ceux de Carlinet *et al.* (2008), comme quoi le fait de visiter

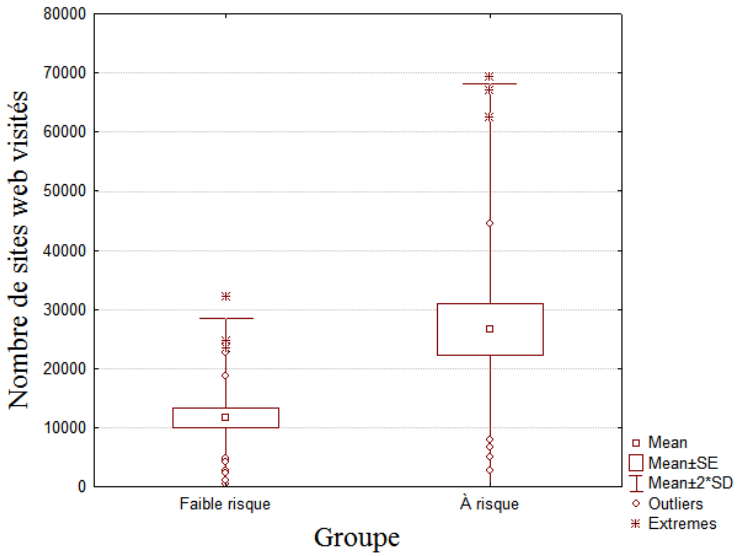


Figure 4.21 Boîtes à moustaches du nombre de sites Web visités

un nombre élevé de sites Web augmenterait les chances de visiter un site malicieux et donc, d'être infecté.

Catégories de sites web visités

Nous avons par la suite désiré analyser si certains types de sites Web exposaient l'usager à un risque d'infection plus élevé. Pour ce faire, nous avons classifié chaque site Web en utilisant l'outil en ligne Safety Center de Trend Micro (2012). Nous avons comptabilisé le nombre de sites Web visités par catégorie afin de déterminer les 10 catégories les plus populaires auprès des participants (Figure 4.22). Au palmarès, nous retrouvons les réseaux sociaux, les engins de recherche ainsi que les sites Web de messagerie courriel.

Nous avons réalisé une régression logistique où la variable indépendante est représentée par le nombre de sites Web visités pour chaque catégorie. Tel que montré dans le Tableau 4.6, nos résultats nous pemettent d'identifier huit catégories à risque : flux vidéo/MP3, pair à pair, infrastructure Internet, téléchargement logiciels, sports, réseaux sociaux, ordinateurs/Internet et jeux de pari.

L'analyse basée sur la régression générale donne des résultats similaires, avec cinq catégories à risque. Des résultats statistiquement significatifs ont été trouvés pour les catégories suivantes : infrastructure Internet, téléchargement logiciels, ordinateurs/Internet, pornographie et traducteur/page cache.

Selon le rapport 2011 de Symantec sur les menaces Internet (Symantec Corporation (2012)), 61% des sites Web malicieux seraient en fait des sites réguliers qui auraient été

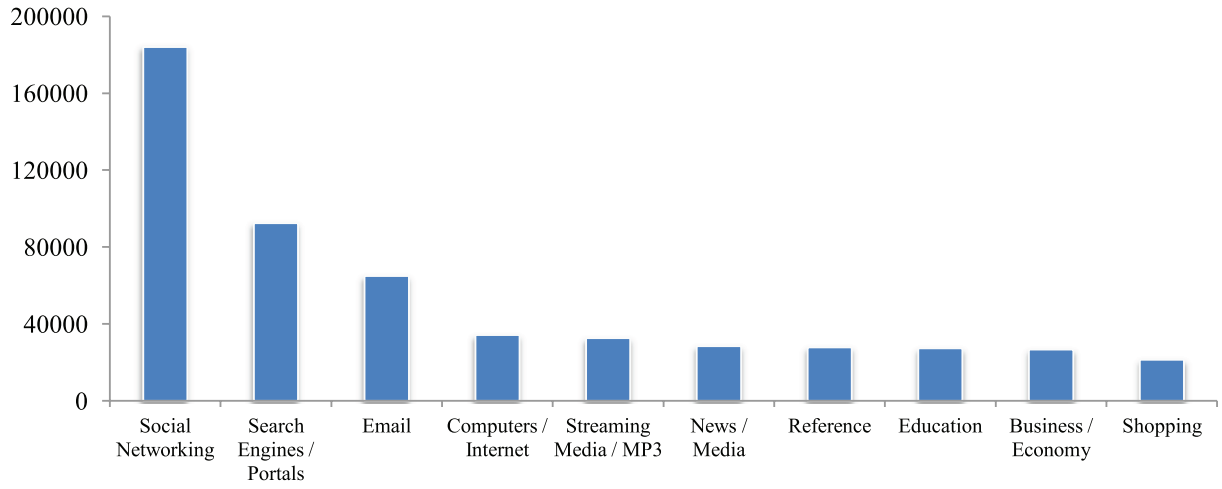


Figure 4.22 Classement des 10 catégories de sites Web les plus populaires

compromis et infectés avec du code malveillant. Plus précisement, les 10 catégories de sites les plus à risque en 2011 étaient : blogue/communications web, hébergement/sites personnel, affaires/économie, magasinage, éducation/référence, technologie et Internet, divertissement et musique, automobile, santé et médecine, pornographie. En comparaison, nos résultats viennent confirmer que les catégories les plus à risque sont pour la grande majorité des sites Web réguliers. De plus, notre analyse vient confirmer en partie les résultats obtenus par Carlinet *et al.* (2008) qui ont, en analysant le trafic réseau de plusieurs milliers d'usagers, trouvé qu'une utilisation élevée de réseaux pair à pair et d'applications de flux vidéo augmente les risques de générer du trafic malveillant.

4.3 Discussion des résultats

Nos résultats démontrent que l'usager influence de par ses caractéristiques et son comportement son exposition aux logiciels malveillants ainsi que son risque d'infection. Qui plus est, le succès ou l'échec d'une majorité grandissante de menaces informatiques repose sur une action (ou une inaction) de l'utilisateur. Il devient ainsi primordial d'impliquer l'usager dans les tests d'antivirus si nous souhaitons évaluer les performances des produits dans un contexte qui soit représentatif de l'expérience des usagers. De meilleures évaluations permettront aux compagnies d'antivirus de comprendre comment leurs produits performent dans un contexte réel, en plus de les aider à mieux comprendre les faiblesses de ces derniers. Ainsi, les compagnies antivirus pourront investir davantages de ressources dans le développement de

nouvelles méthodes de protection au lieu d'investir dans des tests antivirus qui servent plus souvent d'argument de vente qu'à améliorer les produits.

Nous espérons avoir démontré par la présente expérience qu'il est dorénavant essentiel de prendre en compte l'utilisateur lorsque nous voulons évaluer les performances non seulement des antivirus, mais des produits de sécurité en général. À cet effet, nous sommes d'avis que les essais cliniques appliqués en sécurité informatique offrent une solution viable qui permettra de mieux évaluer l'efficacité de différentes solutions de protection en plus d'aider à approfondir notre compréhension des facteurs de risque d'infection. Le recours à de telles expériences peut non seulement permettre d'améliorer les méthodes de protection, mais il peut en plus contribuer à l'amélioration des méthodes de prévention.

Cependant, la réussite de ces méthodes repose sur une connaissance a priori des facteurs de risque. Or, certaines méthodes de prévention sont basées sur des assomptions qui ne reflètent pas nécessairement la réalité, comme le fait que les personnes âgées seraient plus à risque ou encore qu'un niveau élevé d'expertise en informatique serait un facteur protectif. Bien que nous avons présenté dans ce mémoire plusieurs travaux qui visent à identifier les facteurs de risque d'infection, la majorité de ces études sont basées sur des questionnaires et non sur des données réelles. Il devient dès lors nécessaire de réaliser des expériences dites de terrain avec de vrais usagers si nous désirons valider ces assomptions afin de mieux orienter nos efforts en terme de formation et de sensibilisation des usagers.

4.3.1 Limitations des résultats

Les résultats que nous avons obtenus sont sujets à certaines limitations. Le calcul de l'efficacité de l'antivirus est limité aux 113 menaces détectées et il ne prend pas en considération la présence de faux positif dans l'échantillon. De plus, le nombre de menaces manquées par l'antivirus est possiblement sous-estimé. Seulement les infections détectées par notre protocole lors des rencontres mensuelles ont été prises en compte.

Compte tenu du fait que nous avons effectué le recrutement sur un campus universitaire, notre bassin de participants potentiels était principalement composé de jeunes étudiants. Par conséquent, les jeunes ainsi que les étudiants sont surreprésentés dans notre échantillon. De plus, un possible biais géographique a été introduit puisque la grande majorité des participants résident dans la région du Grand Montréal. Ces différentes limitations au niveau de l'échantillon des usagers peuvent expliquer entre autres pourquoi nous n'avons trouvé qu'aucun facteur démographique présente un risque d'infection plus élevé. Bien que la taille de la population soit limitée comparativement aux expériences dans le milieu médical, notre étude est comparable aux études de terrain réalisées dans le domaine de l'utilisabilité.

Une autre biais potentiel provient de l'expérience elle-même. Sachant qu'ils participaient

à une expérience, certains usagers ont peut-être modifié leur utilisation de l'ordinateur. Nous avons posé la question aux participants lors de leur dernière rencontre et 43 ont affirmé ne pas avoir modifié leur comportement. Parmi les 7 autres usagers, 2 ont mentionné avoir modifié leur comportement afin de respecter les conditions de l'expérience (ne pas réinstaller le système d'exploitation, créer de partition, etc.), 2 ont spécifié qu'ils n'ont volontairement pas effectué d'activités gênantes sur l'ordinateur, un usager a mentionné qu'il n'a pas visité de sites Web bancaires sécurisés, un autre a dit s'être forcé à utiliser le plus souvent l'ordinateur, alors que le dernier a mentionné avoir contrôlé l'accès à son ordinateur afin de s'assurer d'être le seul utilisateur.

Bien que nous ayons été en mesure de déterminer plusieurs facteurs de risque par corrélation, il se peut que certains des résultats aient été faussement déclarés statistiquement significatifs. Ce problème particulier apparaît lorsque plusieurs tests statistiques sont réalisés sur les mêmes données, ce qui augmente la probabilité de déclarer des résultats faussement significatifs. Il existe toutefois plusieurs méthodes de correction afin d'ajuster les valeurs de p et de diminuer le nombre de faux positif. Parmi les plus répandues, nous retrouvons, entre autres, la technique de Holm et la correction de Bonferroni, qui implique de multiplier la valeur p par le nombre de tests effectués. Toutefois, ces méthodes ne sont qu'optimales lorsque les tests réalisés ne sont pas corrélés. Une analyse supplémentaire des données devra donc être réalisée afin de déterminer la méthode d'ajustement la plus appropriée dans notre cas.

Finalement, les facteurs de risque identifiés ne peuvent suffir à expliquer le lien de cause à effet avec les infections. À cet effet, une analyse détaillée de la source et du moyen d'infection pour chacune des 113 détections est en cours. Ce n'est qu'une fois ces résultats connus que nous serons en mesure de valider l'impact réel des facteurs de risque d'infection précédemment identifiés.

4.4 Conclusion

Nous avons dans ce chapitre présenté nos résultats reliés à l'évaluation de l'antivirus étudié et à l'impact des facteurs humains sur le risque d'infection.

La première partie de l'analyse, soit celle qui vise à évaluer l'antivirus, nous a permis de constater que 38% des participants ont été exposés à des menaces qui ont été bloquées par l'antivirus. De plus, 20% des participants ont été exposés à des menaces qui ont échappées à l'antivirus malgré le fait que ce dernier était mis à jour quotidiennement. Nous avons été en mesure de calculer que le logiciel antivirus étudié offre une efficacité de 84,07% en terme de détection. Ce résultat est bien en deça de ceux obtenus par la majorité des tests, où il est rare d'obtenir des résultats inférieurs à 90%. L'analyse des questionnaires nous a permis

de mieux comprendre la perception qu'ont les utilisateurs du logiciel antivirus. Bien que nos résultats ne peuvent être généralisés, ils nous ont permis de constater que les utilisateurs ont globalement confiance au niveau de protection fourni par l'antivirus et qu'ils apprécient recevoir des informations de ce dernier, que ce soit concernant les détections ou encore les mises à jour.

La deuxième partie nous a permis d'identifier différents facteurs de risque à l'égard de la probabilité d'infection ainsi que du nombre d'occurence d'infections. Au niveau des facteurs démographiques et des caractéristiques, nous avons obtenu que le genre, l'âge, le statut et le domaine d'activité n'influencent ni la probabilité d'infection ni le nombre de détections. Toutefois, nous avons trouvé que les utilisateurs ayant un niveau d'expertise élevé en informatique présentent un risque plus élevé d'infection. En ce qui concerne le comportement usager, nous avons identifié quatre facteurs de risque qui influencent significativement la probabilité d'infection d'un système : le temps de connexion, le nombre d'hôtes contactés, le nombre de sites Web visités et cinq catégories de sites Web (flux vidéo/MP3, pair à pair, infrastructure Internet, téléchargement logiciels, sports, réseaux sociaux, ordinateurs/Internet et jeux de pari). Nous avons de plus été en mesure d'identifier trois facteurs de risque significatifs qui affectent le nombre de détections des usagers, soit le nombre d'applications installées, le nombre d'hôtes contactés ainsi que cinq catégories de sites Web (infrastructure Internet, téléchargement logiciels, ordinateurs/Internet, pornographie et traducteur/page cache).

CHAPITRE 5

TRAVAUX ET APPLICATIONS FUTURS

Nous exposons dans ce chapitre comment la méthodologie utilisée peut être adaptée afin de réaliser des tests d'antivirus à grande échelle. Nous analysons deux modèles d'évaluation différents en comparant leurs coûts ainsi que leurs avantages et inconvénients respectifs. De plus, nous expliquons comment la méthodologie peut être utilisée afin de réaliser des tests comparatifs d'antivirus. Finalement, nous présentons comment les essais cliniques peuvent permettre d'évaluer différents outils et stratégies appliqués en sécurité informatique.

5.1 Test individuel d'antivirus

La réalisation d'essais cliniques appliqués aux tests individuels d'antivirus pourra permettre aux compagnies de mieux comprendre comment leurs produits performent dans un contexte réel d'utilisation en plus d'identifier les aspects de leurs produits qui peuvent être améliorées, tels que l'interface utilisateur, la détection et la désinfection. Elle pourra par ailleurs aider à comprendre les comportements utilisateur qui entraînent un risque plus élevé d'infection, ce qui entraînera une meilleure compréhension des facteurs de risque tout en permettant de mieux orienter les efforts en terme de formation et de sensibilisation.

Il convient cependant d'adapter la méthodologie de l'étude précédemment réalisée afin de pouvoir mener des tests à grande échelle qui fourniront des résultats significatifs. À cet effet, nous proposons dans les sections qui suivent deux différents modèles qui peuvent être utilisés afin de mener des tests individuels d'antivirus à grande échelle. Le premier modèle est similaire à l'expérience réalisée dans le cadre du présent travail, en ce sens qu'il implique que les participants se déplacent pour assister à des rencontres mensuelles alors que le second modèle est basé sur des accès distants afin d'obtenir les données.

Bien que les deux modèles soient différents, ils partagent plusieurs caractéristiques communes.

Durée : Quatre mois.

Population : Minimum 100 participants utilisant Windows 7 ou Windows 8 sur leur propre ordinateur avec une connexion Internet.

Échantillon : Participants recrutés en fonction de leur profil démographique (âge, genre, statut) de sorte à obtenir un échantillon représentatif. Comme pour l'expérience précédemment réalisée, les utilisateurs devront consentir à l'installation d'outils sur leur

ordinateur afin de permettre la collecte des données. Ils ne devront pas remplacer le système d'exploitation, installer un autre produit antivirus ou supprimer nos outils. Un service de support technique leur sera offert et ils auront la possibilité de se retirer de l'étude en tout temps.

Produit : Un produit qui se verra assigner un minimum de 100 utilisateurs.

Équipement : Contrairement à l'étude pilote que nous avons réalisée, aucun équipement spécifique ne sera acheté puisque les participants devront utiliser leur ordinateur personnel. Toutefois, un minimum de spécifications techniques seront exigées au niveau de la mémoire, du disque dur et du processeur.

5.1.1 Scénario de test avec rencontres

Dans ce premier scénario, les utilisateurs doivent se présenter à un total de 5 rencontres d'environ une heure chaque : une rencontre initiale et quatre rencontres mensuelles. La rencontre initiale sert principalement à installer l'antivirus à évaluer ainsi que des scripts, des applications et des outils de diagnostique. L'utilisateur doit signer le formulaire de consentement en plus de répondre à un questionnaire initial. Au cours des rencontres mensuelles, les participants doivent remplir un questionnaire en ligne pendant que des données sont collectées sur leur ordinateur. Les outils diagnostique sont exécutés afin de déterminer si ce dernier est infecté. Dans l'éventualité où le logiciel antivirus détecte un fichier malicieux ou si nous soupçonnons que l'ordinateur est infecté, un consentement supplémentaire est demandé au participant afin de collecter des données additionnelles qui permettront d'identifier la source et le moyen d'infection.

Dépenses initiales : Étant donné que ce modèle est une adaption de l'expérience précédente, les seules dépenses initiales sont reliées à l'amélioration des scripts et des outils. Si nous considérons un mois de travail de la part d'un assistant technique, le coût ne devrait pas dépasser 5 000$.

Dépenses de fonctionnement : Du travail technique sera nécessaire tout au long de l'expérience. Nous estimons qu'il faudra 5 heures par utilisateur pour l'ensemble des rencontres, une heure par utilisateur pour le support technique et une autre heure par utilisateur pour les tâches administratives. Considérant un total de 2 800 heures pour 100 participants à un taux horaire de 20$, nous obtenons une dépense de 14 000$.

Indemnité compensatoire : Chaque utilisateur recevra gratuitement une license d'un an du produit antivirus installé sur son ordinateur. Nous planifions que ces coûts seront directement couverts par la compagnie antivirus. Chaque participant recevra en plus une compensation de 20$ pour chaque rencontre, pour un total de 100$ par utilisateur.

Avec une expérience impliquant 100 utilisateurs, le coût total ne devrait pas dépasser 10 000$.

Le coût final pour une telle expérience impliquant 100 utilisateurs sur une période de 4 mois devrait se situer aux environs de 29 000$. Ce modèle a comme principal avantage d'être moins dispendieux que l'expérience précédemment réalisée puisqu'il n'implique pas l'achat d'ordinateurs pour l'ensemble des participants. De plus, il permet aux chercheurs d'avoir un accès direct aux ordinateurs. Toutefois, il présente deux inconvénients majeurs. Premièrement, ce type d'expérience demande énormément de temps et d'efforts puisqu'il nécessite de rencontrer chaque utilisateur et d'inspecter leur ordinateur. Deuxièmement, il introduit un biais géographique compte tenu du fait que les utilisateurs doivent résider à proximité de l'endroit où se déroule l'étude.

5.1.2 Scénario de test avec accès distant

Le second scénario est entièrement réalisé à distance. En début d'expérience, les utilisateurs sont redirigés vers un site Web sécurisé où ils peuvent télécharger la solution logicielle nécessaire au fonctionnement de l'expérience. Lors de son exécution, la solution supprime tout logiciel antivirus installé avant d'installer le produit antivirus qui sera sujet à évaluation ainsi que les scripts, applications et outils de diagnostique requis. Une fois l'installation complétée, les scripts envoient à nos serveurs dédiés une image de l'état actuel du système afin de servir de référence. Chaque mois, l'utilisateur doit remplir un questionnaire en ligne alors que nos scripts envoient à nos serveurs les données compilées au cours du dernier mois. Bien que l'étape d'envoi des données soit automatique, du travail technique sera nécessaire afin d'analyser l'état du système de sorte à déterminer si ce dernier est infecté ou soupçonné de l'être. Dans un tel cas, l'utilisateur est dirigé vers un second formulaire de consentement afin de nous autoriser à collecter des informations additionnelles qui nous permettront de déterminer la source et le moyen d'infection. Cette collecte de données serait soit déclenchée automatiquement par l'utilisateur lui-même ou par un accès distant à l'ordinateur. Étant donné que les informations seront envoyées à distance à nos serveurs, elles devront être chiffrées afin de préserver l'identité et la sécurité de l'utilisateur, mais aussi de nos infrastructures.

Dépenses initiales : Compte tenu du fait que ce scénario présente plus de défis techniques, nous prévoyons au minimum 4 mois de travail afin de développer et de tester la solution logicielle, les outils et les scripts. Les dépenses initiales devraient donc être de l'ordre de 20 000$.

Dépenses de fonctionnement : Du travail technique sera aussi nécessaire afin d'assurer le bon fonctionnement de l'expérience : deux heures par utilisateur pour l'analyse men-

suelle des données, une heure par utilisateur pour le support technique et une heure par utilisateur pour les tâches administratives. Avec un total de 1 600 heures pour 100 participants considérant un taux horaire de 20$, le coût total devrait être d'environ 8 000$.

Indemnité compensatoire : Les participants se verront offrir une license d'un an du logiciel antivirus installé sur leur ordinateur. Nous prévoyons que ces coûts seront assumés par la compagnie antivirus. Les utilisateurs recevront de plus une compensation sous la forme d'un certificat cadeau d'une valeur de 50$ s'ils complètent la durée totale de l'expérience. Avec 100 participants, le total ne devrait pas excéder 5 000$.

Les dépenses reliées à une expérience de 4 mois basée sur ce scénario avec 100 participants devraient être de l'ordre de 33 000$. Ce modèle d'expérience est donc légèrement plus dispendieux que le modèle précédent puisqu'il nécessite un effort de développement supplémentaire. Bien que cette approche ne nous donne pas un accès direct à l'ordinateur, elle offre l'avantage de permettre un recrutement à l'international puisque les participants ne seront pas contraints d'assister à des rencontres.

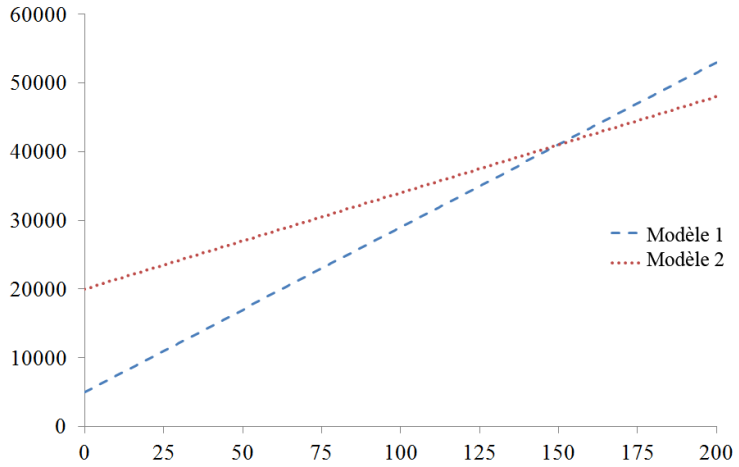


Figure 5.1 Dépenses en dollars en fonction du nombre de participants

La Figure 5.1 montre une comparaison des dépenses pour chaque scénario en fonction du nombre d'utilisateurs pour une période de 4 mois. Alors que le premier modèle présente un coût moins élevé pour une expérience avec 100 participants, ce n'est pas le cas lorsque le nombre de participants devient supérieur à 150. Si nous négligeons les dépenses initiales qui s'appliqueront uniquement lors d'une première expérience, le second modèle apparaît comme celui étant le plus avantageux pour la réalisation d'expériences impliquant plus de 150 participants. En d'autres mots, le premier scénario est plus rapide en terme de développement,

mais le second scénario semble plus avantageux dans une perspective de commercialisation à grande échelle.

5.2 Test comparatif d'antivirus

Une large portion des tests d'antivirus est destinée à identifier quel produit performe mieux qu'un autre, que ce soit pour permettre aux utilisateurs de faire un choix plus éclairé lors de leur achat ou encore pour aider les fournisseurs d'antivirus à déterminer leurs stratégies de recherche et développement et de marketing. Compte tenu de l'avantage en terme de réalisme et de disponibilité des données utilisateurs de la méthodologie utilisée, la question qui s'en suit est s'il est possible d'adapter cette méthodologie afin de réaliser des tests comparatifs d'antivirus.

Une des problématiques des tests comparatifs réalisés en laboratoire consiste à s'assurer que tous les produits antivirus sont évalués sous exactement les mêmes conditions. Non seulement ils doivent être testés dans le même environnement, mais ils doivent être exposés aux mêmes menaces en même temps. Alors qu'il est relativement facile de guarantir ces conditions lorsque les tests sont réalisés dans un environnement contrôlé, la tâche n'est pas aussi simple dans le cas d'une étude utilisateur réalisée sur le terrain. L'exposition des produits antivirus aux menaces ne peut être contrôlée puisqu'elle dépend des utilisateurs. Toutefois, ce type d'étude a l'avantage de ne présenter aucun biais quant au choix des menaces puisque ces dernières sont « choisies » par les utilisateurs qui n'ont a priori aucun parti pris dans le processus d'évaluation. Une importante population d'utilisateurs permettrait de garantir l'exposition de chaque produit évalué à un large échantillon de menaces de telle sorte à obtenir des résultats statistiquement significatifs. En d'autres mots, le recours aux études utilisateurs comme méthode d'évaluation comparative doit impliquer un nombre élevé d'utilisateurs durant une longue période de temps afin d'éliminer tout biais au niveau de la sélection des menaces. À cet effet, nous postulons qu'une telle étude utilisateur devrait inclure au moins 100 participants par produit antivirus évalué et durer au minimum 4 mois afin de garantir l'obtention de résultats statistiquement significatifs.

5.3 Autres applications à la sécurité informatique

Le recours aux essais cliniques peut être étendu à l'évaluation de divers stratégies et produits de sécurité. De telles expériences peuvent contribuer à améliorer notre compréhension des mécanismes d'infection et des différents facteurs de risque. Une meilleure connaissance des comportements à risque permettra de mieux orienter les efforts en terme de formation et de sensibilisation. Les individus identifiés comme les plus à risque pourront bénéficier d'un

programme de sécurité sur mesure, adapté à leurs besoins. Ce programme pourrait prendre la forme de différentes politiques de sécurité, de produits de sécurité additionnels ou encore d'une formation particulière, permettant aux entreprises de concentrer leurs ressources là où le risque est le plus élevé (Lee (2012)).

5.4 Conclusion

Nous avons présenté comment la méthodologie utilisée peut être adaptée à la réalisation de tests antivirus à grande échelle, tant individuels que comparatifs. À cet effet, nous avons proposé deux modèles différents d'expériences et comparé leurs coûts associés. Finalement, nous avons démontré comment la méthode des essais cliniques peut être appliquée à d'autres aspects de la sécurité informatique.

CHAPITRE 6

CONCLUSION

L'augmentation de la complexité tant au niveau des logiciels malveillants que des produits antivirus a rendu difficile l'évaluation adéquate de ces derniers. Alors que la majorité des tests d'antivirus sont réalisés en laboratoire, ils ne peuvent prendre en compte une multitude de facteurs. Par conséquent, ces tests ne peuvent refléter les performances réelles des produits. Une meilleure méthode de test est dès lors nécessaire afin d'aider l'industrie à développer de meilleurs produits de sécurité, et à aider l'utilisateur à faire un choix de produit éclairé qui réponde à ses besoins.

Nous avons, dans le cadre du présent mémoire, proposé une méthode d'évaluation inspirée des essais cliniques utilisés dans le domaine médical. Une étude utilisateur impliquant la participation de 50 usagers durant quatre mois a été réalisée afin d'évaluer les performances d'un produit antivirus, ce qui constitue, au meilleur de nos connaissances, une première dans le domaine de la sécurité informatique. Au terme de cette dernière, nous avons été en mesure d'atteindre nos objectifs de recherche, soit d'évaluer un logiciel antivirus, de déterminer les facteurs de risque d'infection et de valider la viabilité de la méthodologie proposée.

Notre expérience nous a permis d'appliquer la méthodologie des essais cliniques au contexte de la sécurité informatique. Plus précisement, nous avons été en mesure de calculer la performance en terme de détection d'un produit antivirus et d'évaluer la perception qu'en ont les utilisateurs. En plus de fournir des résultats plus réalistes, nous croyons que cette nouvelle méthode de test reflète davantage les besoins réels des utilisateurs.

Quant aux facteurs de risque, nos analyses ont démontré que certaines caractéristiques et comportements usagers sont associés à un risque d'infection plus élevé. Au niveau des caractéristiques, nous avons trouvé que les utilisateurs ayant un niveau d'expertise élevé en informatique sont exposés à un risque plus élevé d'infection alors que l'âge, le genre, le statut et le domaine d'activité n'auraient pas d'impact significatif. En ce qui concerne le comportement usager, nous avons identifié plusieurs facteurs de risque tels que le nombre d'applications installées, le temps de connexion, le nombre d'hôtes contactés et le nombre de sites Web visités. À cet égard, nous avons constaté que plusieurs catégories de sites Web légitimes présentent des risques plus élevés pour l'utilisateur comme par exemple les sites de flux vidéo/MP3, les sites de téléchargement logiciels ou encore les sites d'ordinateurs/Internet.

Bien que certains de ces facteurs de risque peuvent sembler intuitifs, leur impact respectif n'avait jamais été confirmé par une expérience scientifique. Une meilleure connaissance

des comportements les plus à risque permettra non seulement de développer de meilleures formations et stratégies de sécurité, mais de valider les solutions actuellement utilisées.

Cette étude pilote basée sur la méthodologie des essais cliniques a démontré que cette approche constitue une alternative viable aux tests réalisés en laboratoire. Cette méthode offre des résultats qui sont de loin plus significatifs et moins sujets à la controverse compte tenu du réalisme de l'environnement de test et de l'indépendance dans la sélection des menaces. Toutefois, nous croyons que les essais cliniques ne peuvent remplacer les tests réalisés en laboratoire, mais qu'ils doivent au contraire être utilisés en complément.

Le présent travail illustre les mérites de réaliser des essais cliniques à grande échelle afin d'évaluer les antivirus et d'autres produits de sécurité. Nous avons à cet effet présenté deux modèles d'expérience. Le premier, qui implique de rencontrer les utilisateurs, devrait être considéré comme une transition vers le second, qui utilise des accès distants pour collecter les données. Ce second modèle présente plusieurs défis techniques et requiert un niveau d'investissement supérieur afin d'être mis en place. Cependant, il n'implique pas de rencontrer les participants, ce qui permet d'élargir le recrutement à l'international. Par conséquent, nous croyons que ce modèle représente une solution plus viable et intéressante dans une persceptive à long terme, tant pour l'industrie antivirus que pour la communauté scientifique.

Les travaux futurs relatifs à notre projet consistent dans un premier temps à poursuivre l'évaluation de l'antivirus. Une analyse supplémentaire sera effectuée afin d'identifier d'éventuelles infections qui auraient échappé à l'antivirus et à notre procole. Nous avons de plus l'intention de compléter l'évaluation de l'antivirus par les usagers en poursuivant l'analyse des questionnaires.

Dans un second temps, nous allons poursuivre l'évaluation des facteurs de risque en analysant non seulement le comportement des utilisateurs à partir des données collectées, mais le comportement rapporté par l'entremise des questionnaires usagers. La connaissance de ces facteurs pourra nous permettre entre autres de développer un modèle mathématique plus précis visant à prédire la probabilité d'infection d'un usager en fonction de ses caractéristiques et son comportement.

Finalement, nous avons l'intention de réaliser une analyse détaillée des 113 détections afin de déterminer les sources et les moyens d'infections utilisés pour compromettre les ordinateurs. Ces résultats permettront d'obtenir des statistiques réelles sur les causes d'infections, en plus de confirmer le rôle de l'utilisateur dans les mécanismes d'infections.

À plus long terme, nous espérons pouvoir continuer notre recherche et réaliser dans le futur des expériences à plus grande échelle afin de mieux évaluer les mérites relatifs de différentes stratégies et solutions appliquées à la sécurité informatique.

RÉFÉRENCES

ANTI-MALWARE TESTING STANDARDS ORGANIZATION (2008a). Best practices for dynamic testing. http://www.amtso.org/released/20081031_AMTSO_Best_Practices_For_Dynamic_Testing.pdf.

ANTI-MALWARE TESTING STANDARDS ORGANIZATION (2008b). The fundamental principles of testing. http://www.amtso.org/released/20081031_AMTSO_Fundamental_Principles_of_Testing.pdf.

AV COMPARATIVES (2013a). File detection test of malicious software. Rapport technique, AV Comparatives.

AV COMPARATIVES (2013b). It security survey 2013. Rapport technique, AV Comparatives.

BASHARI RAD, B., MASROM, M. et IBRAHIM, S. (2011). Evolution of computer virus concealment and anti-virus techniques : A short survey.

CARLINET, Y., ME, L., DEBAR, H. et GOURHANT, Y. (2008). Analysis of computer infection risk factors based on customer network usage. *Second International Conference on Emerging Security Information, Systems and Technologies (SECURWARE'08)*. 317–325.

CLEMENTI, A. (2007). Anti-virus testing websites. Rapport technique, AV Comparatives. http://www.av-comparatives.org/seiten/ergebnisse/AVTW.pdf.

CONTU, R. et CHEUNG, M. (2013). Market Share Analysis : Security Software, Worldwide, 2012. http://www.gartner.com/id=2486015.

DE LUCA, A., LANGHEINRICH, M. et HUSSMANN, H. (2010). Towards understanding ATM security : a field study of real world ATM use. *ACM Symposium on Usable Privacy and Security*.

DUNLAP, C. (2009). False sense of security : New anti-virus testing methodologies are critical to educate customers. http://www.trendmicro.tw/cloud-content/tw/pdfs/about/wp_dunlap_need-for-new-testing-methods.pdf.

ESET (2009). The curious act of anti-malware testing. http://www.eset.com/us/resources/white-papers/Curious_Act_Of_Anti_Malware_Testing.pdf.

FAULHABER, J. et FELSTEAD, D. (2011). Zeroing in on malware propagations methods. Rapport technique, Microsoft Security Intelligence Report.

GORDON, S. et FORD, R. (1996). Real world anti-virus product reviews and evaluations : the current state of affairs. *National Information Systems Security Conference*.

HALLORAN, E., LONGINI, I. M. et STRUCHINER, C. J. (1999). Design and interpretation of vaccine field studies. *Epidemiological Reviews*, 21, 73–88.

HANSMAN, S. et HUNT, R. (2003). A taxonomy of network and computer attack methodologies. Rapport technique, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand.

HARLEY, D. (2009). Making sense of anti-malware comparative testing. *Information security technical report*, 7–15.

HELENIUS, M. (2002). *A System to Support the Analysis of Antivirus Products' Virus Detection Capabilities*. Thèse de doctorat, University of Tampere.

KOSINAR, P., MALCHO, J., MARKO, R., et HARLEY, D. (2010). AV testing exposed. *20th Virus Bulletin International Conference*.

LALONDE LÉVESQUE, F., DAVIS, C., FERNANDEZ, J., CHIASSON, S. et SOMAYAJI, A. (2012a). Methodology for a field study of anti-malware software. *Financial Cryptography and Data Security - FC 2012 Workshops (USEC12)*. 80–85.

LALONDE LÉVESQUE, F., DAVIS, C., FERNANDEZ, J. et SOMAYAJI, A. (2012b). Evaluating antivirus products with field studies. *Proceedings of the 2013 Virus Bulletin International Conference*. 87–94.

LALONDE LÉVESQUE, F., NSIEMPBA, J., FERNANDEZ, J., CHIASSON, S. et SOMAYAJI, A. (2013). Should grandma be on the internet ? a clinical study of risk factors in malware. *Proceedings of the 2013 ACM conference on Computer and communications security*.

LEE, M. (2012). Who's next ? identifying risk factors for subjects of targeted attacks. *Proceedings of the 22th Virus Bulletin International Conference*.

MCAFEE (2012). Consumer Alert : McAfee releases results of global unprotected rates study. `http://web.archive.org/web/20130121021450/https://blogs.mcafee.com/consumer/family-safety/mcafee-releases-results-of-global-unprotected-rates`.

MILNE, G. R., LABRECQUE, L. I. et CROMER, C. (2009). Toward an understanding of the online consumer's risky behavior and protection practices. *Journal of Consumer Affairs*, 43, 449–473.

MORALES, J. A., SANDHU, R. et S.XU (2010). Evaluating detection and treatment effectiveness of commercial anti-malware programs". *Proceedings of the 5th International Conference on Malicious and Unwanted Software (MALWARE)*.

NGO, F. T. et PATERNOSTER, R. (2011). Cybercrime victimization : An examination of individual and situational level factors. *International Journal of Cyber Criminology*, 5, 773–793.

PANDA SECURITY LABS (2012). PandaLabs Quaterly Report January-March 2012. http://press.pandasecurity.com/wp-content/uploads/2012/05/Quarterly-Report-PandaLabs-January-March-2012.pdf.

PC SECURITY LABS (2013). Security solution review on windows 8 platform. Rapport technique, PC Security Labs.

RODE, J. A. (2009). Digital parenting : designing children's safety. *British Human Computer Interaction Conference (British HCI)*. 244–251.

SHENG, S., HOLBROOK, M., KUMARAGURU, P., CRANOR, L. F. et DOWNS, J. (2010). Who falls for phish ? A demographic analysis of phishing susceptibility and effectiveness of interventions. *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 373–382.

SOLIC, K. et ILAKOVAC, V. (2009). Security perception of a portable PC user (the difference be-tween medical doctors and engineers) : A pilot study. *Medicinski Glasnik*. vol. 6, 261–264.

SOMAYAJI, A., LI, Y., INOUE, H., FERNANDEZ, J. M. et FORD, R. (2009). Evaluating security products with clinical trials. *Workshop on Cyber Security Experimentation and Test (CSET)*.

SURFRIGHT (2009). 32% of Computers Still Infected, Despite Presence of Antivirus Program. http://www.surfright.nl/en/home/press/32-percent-infected-despite-antivirus.

SYMANTEC CORPORATION (2012). Internet security threat report 2011 trends. Rapport technique 18, Symantec Corporation.

SZOR, P. (2005). *The Art of Computer Virus Research and Defense*. Addison-Wesley Professional.

TREND MICRO (2012). Website classification. http://solutionfile.trendmicro.com/solutionfile/Consumer/new-web-classification.html.

VON NEUMANN, J. (1951). The general and logical theory of automata. *Cerebral mechanisms in behavior*, 1–41.

VRABEC, J. et HARLEY, D. (2010). Real performance ? *EICAR Annual Conference*.

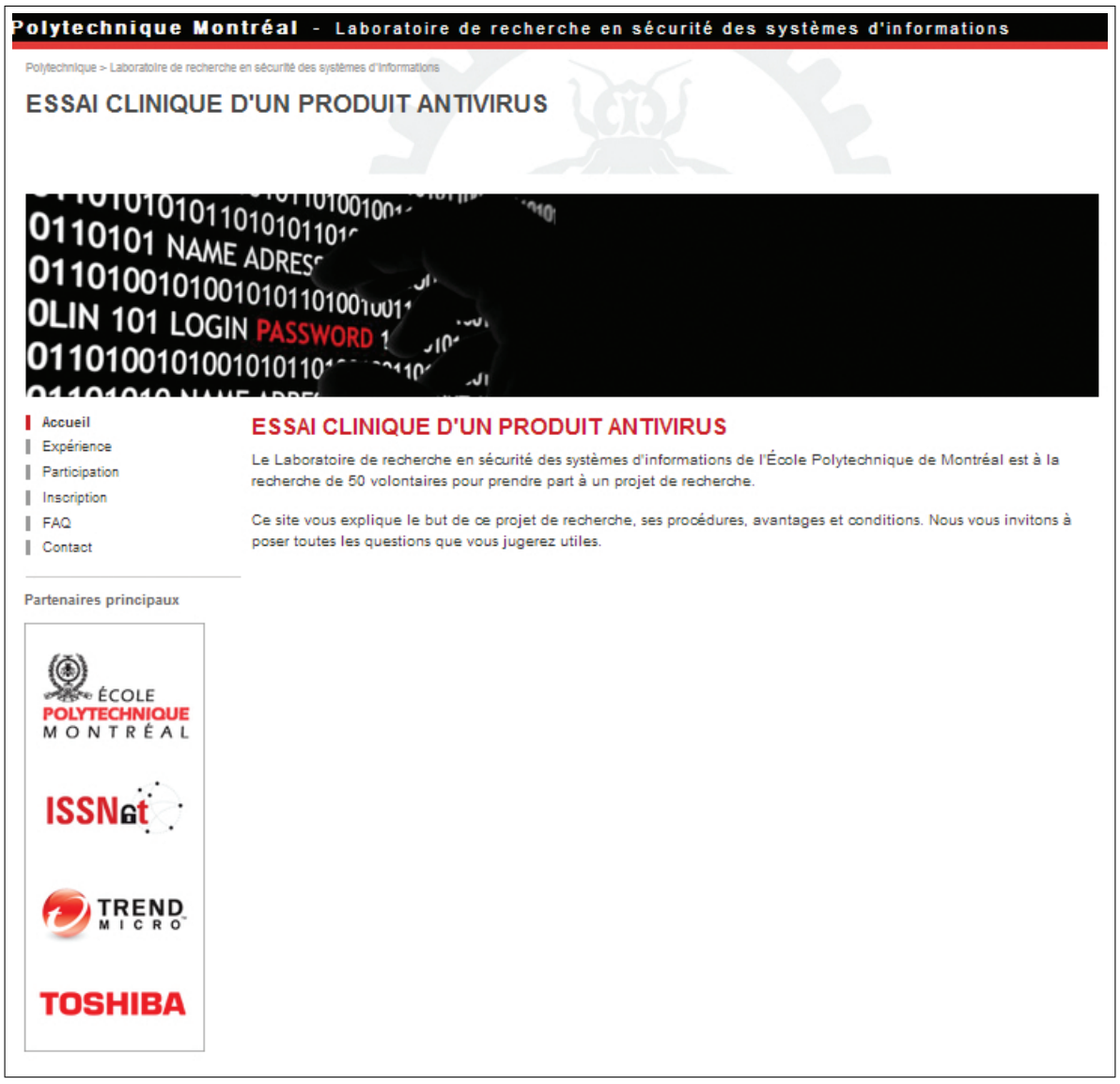
WASH, R. (2010). Folk models of home computer security. *ACM Symposium on Usable Privacy and Security (SOUPS)*.

ANNEXE A

AFFICHE DE RECRUTEMENT

ANNEXE B

SITE WEB



ANNEXE C

FORMULAIRE DE CONSENTEMENT

Formulaire d'information et de consentement

Titre du projet de recherche
Essai clinique de produit antivirus

Personnel de recherche
Ce projet de recherche est mené par des chercheurs du *Laboratoire de recherche en sécurité des systèmes d'information*s (SecSI) du Département de Génie Informatique et de Génie Logiciel à l'École Polytechnique de Montréal.

Chercheur responsable
José M. Fernandez, ing. PhD
Professeur adjoint
Responsable du laboratoire SecSI

Co-chercheur
Carlton Davis, PhD
Stagiaire Postdoctoral

Étudiant
Fanny Lalonde Lévesque,
Étudiante M.Sc.A. sous la direction du professeur José M. Fernandez
Responsable du soutien technique auprès des sujets de recherche

Responsable technique
Pier-Luc St-Onge, M.Sc.A.
Analyste

Introduction
Nous sollicitons votre participation à un projet de recherche. Cependant, avant d'accepter de participer à ce projet et de signer le formulaire de consentement, veuillez prendre le temps de lire, de comprendre et de considérer attentivement les renseignements qui suivent. Ce document vous explique le but de ce projet de recherche, ses procédures, avantages et risques.

Nous vous invitons à poser toutes les questions que vous jugerez utiles à la personne qui vous présente ce document et à lui demander de vous expliquer tout mot ou renseignement qui n'est pas clair.

Nature et objectifs du projet de recherche
L'étude réalisée a trois principaux objectifs : (1) déterminer quels sont les facteurs qui influencent le nombre d'infections d'un système; (2) évaluer la performance d'un logiciel antivirus; et (3) identifier les vecteurs d'infection et les moyens par lesquels un système se retrouve infecté.

La participation au projet de recherche est réservée à des personnes majeures et nécessite le recrutement de 50 sujets de recherche.

Déroulement du projet de recherche
La durée totale de l'expérience est de quatre mois. En début de projet, un portable de qualité vous sera vendu au prix de 350$, toutes taxes incluses. Ce dernier contiendra le logiciel antivirus commercial OfficeScan de Trend Micro™ et des outils diagnostiques pour déterminer si votre portable est infecté. Veuillez noter que vous serez libres de personnaliser votre portable en installant les applications de votre choix.

À chaque fin de mois, vous devrez vous présenter à l'École Polytechnique de Montréal pour une courte rencontre dont la durée pourra varier entre 1h à 2h. Durant ces rencontres mensuelles, il vous sera demandé de remplir un formulaire en ligne. Ce formulaire nous permettra de mieux évaluer votre expérience à l'égard du logiciel antivirus installé et de recueillir des informations générales concernant votre utilisation du portable. Ces rencontres nous permettrons aussi de recueillir les différentes informations compilées sur votre portable et de déterminer si ce dernier a été infecté. Si le logiciel antivirus trouve une infection sur votre portable ou si nos outils diagnostiques nous suggèrent que votre portable est infecté, nous vous demanderons une autorisation spéciale afin de recueillir des données supplémentaires nécessaires afin de déterminer le moyen et la source de l'infection. Dans l'éventualité où le logiciel antivirus ne serait pas en mesure de supprimer une infection, nous communiquerons les informations nécessaires à Trend Micro™ afin qu'ils puissent développer une méthode de détection et de désinfection qu'ils pourront inclure dans les mises à jour du produit antivirus.

Collaboration du sujet au projet de recherche

- Ne pas formater le portable utilisé dans le cadre de l'étude.
- Ne pas remplacer le système d'exploitation installé sur le portable.
- Ne pas créer de partition(s) sur le portable.
- Ne pas installer un autre produit antivirus sur le portable.
- Ne pas supprimer les outils diagnostiques installés sur le portable.
- Ne pas supprimer les fichiers générés automatiquement sur le portable dans le cadre de l'expérience.
- Contacter en premier le personnel de recherche et non le fournisseur du portable en cas de problème.
- Se présenter à un total de quatre rencontres mensuelles.

Avantages et risque

En participant au projet de recherche, vous bénéficierez des avantages suivants :

- Obtenir un portable de qualité évalué à 439$ par Toshiba pour seulement 350$, toutes taxes incluses.
- Obtenir une protection d'un an par le logiciel antivirus OfficeScan de Trend Micro™.
- Obtenir une compensation financière pouvant aller jusqu'à 400$.
- Obtenir une garantie de 3 ans pour le portable disponible au comptoir CoopHEC.

Cette expérience ne comporte aucun risque additionnel à une utilisation courante du portable.

Participation volontaire et possibilité de retrait

Votre participation à ce projet de recherche est volontaire. Vous êtes donc libre de refuser d'y participer. Vous pouvez également vous retirer de ce projet à n'importe quel moment, sans avoir à donner de raisons, en faisant connaître votre décision au chercheur responsable du projet ou à l'un des membres du personnel affecté au projet. En cas de retrait, une procédure simple vous sera donnée si vous désirez supprimer du portable les outils diagnostiques installés et les informations enregistrées sur votre portable tout au long de l'expérience. Le personnel de recherche sera disponible pour vous guider dans cette procédure, mais vous devrez réaliser vous-même les manipulations. À noter que vous pourrez conserver le portable puisqu'il sera votre propriété.

En cas de perte ou de vol du portable, ou de bris non couvert par la garantie, nous nous réservons le droit de mettre fin à votre participation à l'étude.

Dans l'éventualité de l'arrêt de votre participation à l'expérience, le montant de la compensation financière sera déterminé en fonction du nombre de rencontres auxquelles vous aurez assisté.

À noter que les candidats qui ont l'intention de compléter l'expérience en assistant aux quatre rencontres seront priorisés dans le recrutement.

Confidentialité

Durant votre participation à ce projet, le chercheur responsable ainsi que son personnel recueilleront et consigneront des renseignements vous concernant. Seuls les renseignements nécessaires pour répondre aux objectifs scientifiques de ce projet seront recueillis. La nature des renseignements qui seront recueillis peut être divisée en trois catégories :

- Des données statistiques anonymes concernant votre utilisation du portable, cueillies de façon automatique par les outils installés sur le portable
- Des données brutes concernant une potentielle infection de votre portable, cueillies manuellement par notre équipe, seulement en cas d'infection soupçonnée et seulement si vous en donnez le consentement supplémentaire.
- Des données personnelles vous concernant, cueillies seulement à partir de vos réponses aux questionnaires de participation à l'expérience.

Les données statistiques anonymes recueillies automatiquement sur votre ordinateur comprennent :

- Le nombre et le type d'applications installées sur votre portable.
- Les différents navigateurs web installés sur votre portable ainsi que leurs configurations.
- Le nombre moyen d'heures par jour que votre portable est connecté à Internet.
- Le nombre moyen d'heures par jour que votre portable est allumé.
- Le type et l'adresses IP des réseaux à travers lesquels vous vous connecté à Internet.
- Le nombre de sites web visités.
- Le nombre et le type de fichiers téléchargés à partir d'Internet.

Dans le cas où notre équipe soupçonne une infection du portable, et si vous nous en donnez le consentement, nous procéderons à un examen plus approfondi du portable afin d'essayer de déterminer les causes de l'infection. Pour ceci nous aurons à examiner certaines données brutes concernant votre utilisation du portable. Ces données peuvent comprendre :

- Votre historique de navigation sur Internet (sites Web visités)
- Les journaux (logs) de diagnostique d'outils de sécurité (p.ex. logiciel anti-virus)
- Les données reçues et transmises via le réseau

Parmi les données examinées, seules celles qui seront reliées directement à l'infection seront conservées, p.ex., site Web qui a donné lieu à l'infection, fichiers téléchargés et données transmises par les logiciels malveillants, etc.

Finalement, les données personnelles vous concernant et qui seront recueillies via les questionnaires comprennent :

- Votre nom
- Votre numéro de téléphone
- Votre adresse courriel
- Votre sexe
- Votre groupe d'âge
- Votre domaine d'études ou de travail
- Vos expériences et connaissances en informatique
- Vos habitudes d'utilisation d'Internet et des ordinateurs
- Vos habitudes d'utilisation et votre opinion sur les logiciels antivirus

Tous les renseignements recueillis demeureront strictement confidentiels. Afin de préserver votre identité et la confidentialité des renseignements, vous ne serez identifié que par le numéro de votre portable. Le code reliant votre nom à votre dossier de recherche sera accessible au personnel de recherche impliqué dans le projet tout au long de sa durée. Une fois l'expérience terminée, seul le chercheur responsable aura accès à cette information, et

ce pour une durée limitée de 8 mois, après quoi cette information sera supprimée.

Le chercheur responsable et les membres du personnel utiliseront les données à des fins de recherche dans le but de répondre aux objectifs scientifiques du projet décrits dans le formulaire de consentement. Tous les renseignements recueillis au cours du projet de recherche seront conservés dans la zone de haute sécurité du laboratoire SecSI afin d'en assurer la protection. Seul le chercheur responsable et les membres du personnel auront accès à ces données. Les données brutes et les données personnelles de votre dossier seront conservées pendant 1 an. À la fin de cette période, les mesures nécessaires seront prises afin de supprimer adéquatement les données conservées. Les données statistiques anonymes seront conservées pour une période initiale de 7 ans. Après cette période, une nouvelle approbation valide pour 7 ans devra être obtenue.

Une fois l'expérience terminée, une procédure simple vous sera donnée si vous désirez supprimer du portable les outils diagnostiques installés et les informations enregistrées sur votre portable tout au long de l'expérience. Le personnel de recherche sera disponible pour vous guider dans cette procédure, mais vous devrez réaliser vous-même les manipulations.

Les résultats de l'expérience pourront être publiés ou faire l'objet de discussions scientifiques, mais il ne sera pas possible de vous identifier puisque les données statistiques sont anonymes.

Financement du projet de recherche
Le chercheur responsable et l'établissement ont reçu pour mener à bien ce projet un financement du réseau de recherche stratégique ISSNet du Conseil de recherche en sciences naturelles et génie (CRSNG) et de Trend Micro™, un fabricant de logiciel antivirus.

Compensation
Vous recevrez une somme forfaitaire de 50$ suite à votre présence à chacune des trois premières rencontres. Vous recevrez suite à votre quatrième rencontre une somme de 100$. Si vous vous présentez aux quatre rencontres, vous recevrez une compensation supplémentaire de 150$ à la fin du projet. Votre participation au projet de recherche peut donc vous permettre d'obtenir une compensation financière pouvant aller jusqu'à 400$ si vous respectez les termes du présent contrat.

Identification des personnes-ressources
José M. Fernandez, ing. PhD
Professeur adjoint au Département de Génie Informatique et de Génie Logiciel
Responsable du laboratoire SecSI
Courriel : *jose.fernandez@polymtl.ca*
Tél : (514) 340-4711 p.5433

Carlton Davis, PhD
Stagiaire Postdoctoral au Département de Génie Informatique et de Génie Logiciel
Courriel : *carlton.davis@polymtl.ca*
Tél : (514) 340-5121 p.2063

Fanny Lalonde Lévesque
Étudiante M.Sc.A. au Département de Génie Informatique et de Génie Logiciel
Responsable du soutien technique auprès des sujets recherche
Courriel : *fanny.lalonde-levesque@polymtl.ca*
Tél : (514) 340-4711 p.7172

Pier-Luc St-Onge, M.Sc.A.
Analyste au Département de Génie Informatique et de Génie Logiciel
Responsable technique du projet

Courriel : *pier-luc.st-onge@polymtl.ca*
Tél : (514) 340-4711 p. 5065

En cas de problème technique avec votre portable, veuillez-vous adresser directement à Fanny Lalonde Lévesque.

Surveillance des aspects éthiques du projet de recherche
Le comité d'éthique de la recherche (CÉR) de l'École Polytechnique de Montréal a approuvé ce projet de recherche et en assure le suivi. De plus, il approuvera au préalable toute révision et toute modification apportée au formulaire de consentement et au protocole de recherche.

Pour de plus amples informations, vous pouvez contacter directement le président du CÉR de l'École Polytechnique de Montréal :

Bernard Lapierre
Courriel : *bernard.lapierre@polymtl.ca*
Tél : (514) 340-4711 p. 4567

Remerciement
Votre collaboration au projet de recherche est précieuse et nous vous en remercions.

Consentement
J'ai pris connaissance du formulaire de consentement. Je reconnais qu'on m'a expliqué le projet, qu'on a répondu à mes questions et qu'on m'a laissé le temps voulu pour prendre une décision.

Je consens à participer à ce projet de recherche aux conditions qui y sont énoncées.

Nom et signature du sujet de recherche Date

J'ai expliqué au sujet de recherche les termes du présent formulaire de consentement et j'ai répondu aux questions qu'il m'a posées.

Nom et signature de la personne qui obtient le consentement Date

ANNEXE D

QUESTIONNAIRE INITIAL

Questions initiales

Ces questions sont réservées au questionnaire initial. Elles apparaîtront dans un questionnaire séparé.

1) Entrez le numéro d'identification du laptop: _____

2) De quel sexe êtes-vous? [] homme [] femme [] Je préfère ne pas répondre

3) À quel groupe d'âge appartenez-vous?
   a) 18 à 24 ans
   b) 25 à 30 ans
   c) 31 à 35 ans
   d) 36 à 40 ans
   e) 41 à 45 ans
   f) 46 à 50 ans
   g) Plus de 50 ans

4) Quel énoncé s'applique le mieux à votre situation?
   a) Je suis étudiant au baccalauréat; et je suis dans ma ____ année d'étude
   b) Je suis étudiant à la maîtrise; et je suis dans ma ____ année d'étude
   c) Je suis étudiant au doctorat; et je suis dans ma ____ année d'étude
   d) Autre (spécifiez votre occupation): _____

5) Si vous êtes étudiant ou travailleur, quel est votre domaine d'application?
   a) Sciences appliquées
   b) Informatique
   c) Sciences humaines, Arts et Lettres
   d) Sciences pures et Sciences de la santé
   e) Autre (spécifiez): _____

6) Quel est le langage des pages web que vous visitez le plus fréquemment?
   a) Français
   b) Anglais
   c) Autre (spécifiez) : _____

7) Possédez-vous ou avez-vous déjà utilisé un ordinateur avec Windows 7?
   a) Oui
   b) Non, j'utilise : _____
   c) Je ne sais pas

8) Sélectionnez les tâches que vous avez déjà réalisées; si aucune de ces tâches ne s'applique à votre situation, sélectionnez "Aucune de ces réponses".
   [] J'ai installé ou réinstallé un système d'opération sur un ordinateur

[] J’ai configuré un réseau maison
[] J’ai crée une page web
[] Aucune de ces réponses

9) Quel(s) logiciel(s) antivirus avez-vous déjà utilisé(s)? Sélectionnez tous les choix applicables.
[] Trend Micro
[] Symantec (Norton)
[] McAfee
[] AVG
[] ESET
[] Autre (spécifiez): _____
[] Aucun
[] Je ne sais pas

ANNEXE E

QUESTIONNAIRE MENSUEL

Questionnaire principal

Entrez le numéro d'identification du laptop: _____

Les questions 1-11 visent à déterminer l'expérience et l'opinion de l'utilisateur quant au logiciel antivirus

1) En vous basant sur votre expérience du dernier mois, comment qualifierez-vous le niveau d'interaction requis par le logiciel antivirus installé sur votre ordinateur?

a) Le logiciel requiert mon interaction trop fréquemment
b) Le logiciel ne requiert pas assez souvent mon interaction
c) Le niveau d'interaction requis est adéquat

2) Quel énoncé décrit le mieux votre comportement lorsqu'une fenêtre du logiciel antivirus apparaît?

a) Je la ferme sans la lire
b) Je la lis et je suis ces suggestions
c) Je la lis mais je ne suis pas les suggestions
d) Autre (spécifiez): _____

3) En vous basant sur votre expérience du dernier mois, sur une échelle de 1 à 10, comment évalueriez-vous le niveau d'interférence du logiciel antivirus avec votre utilisation courante de l'ordinateur?

Faible interférence Forte interférence

|____|____|____|____|____|____|____|____|____|____|

1 2 3 4 5 6 7 8 9 10

4) À partir de votre expérience, comment le logiciel antivirus affecte la vitesse de votre ordinateur?

a) Il n'y a pas de différence notable au niveau de la vitesse
b) La vitesse du laptop est diminuée
c) Je ne sais pas

5) À partir de votre expérience du dernier mois, sur une échelle de 1 à 10, comment évalueriez-vous le niveau de protection procuré par le logiciel antivirus installé sur votre ordinateur?

Faible niveau de protection Haut niveau de protection

|____|____|____|____|____|____|____|____|____|____|

1 2 3 4 5 6 7 8 9 10

6) Comment qualifieriez-vous le niveau d'information fournie par le logiciel antivirus?

a) Le logiciel antivirus ne fourni pas assez d'information (par exemple, j'aimerais savoir quand il fait ses mises à jour, etc.)
b) Le logiciel antivirus fourni trop d'information

c) Le logiciel antivirus fourni un niveau adéquat d'information

7) Sur une échelle de 1 à 10, comment qualifieriez-vous le niveau d'utilité des informations fournies par le logiciel antivirus?

Peu utile Très utile

|____|____|____|____|____|____|____|____|____|____|

1 2 3 4 5 6 7 8 9 10

8) Sur une échelle de 1 à 10, comment quantifieriez-vous votre compréhension des informations fournies par le logiciel antivirus?

Difficile à comprendre Facile à comprendre

|____|____|____|____|____|____|____|____|____|____|

1 2 3 4 5 6 7 8 9 10

9) Quel énoncé décrit le mieux votre sentiment lorsqu'une fenêtre du logiciel antivirus apparaît?

[] Je me sens réconforté à l'idée que le logiciel antivirus fasse son travail

[] Je me sens ennuyé que le logiciel antivirus m'interrompe

[] J'ai l'habitude, je ne le remarque plus

[] Autre (spécifiez): _____

10) Avez-vous vu apparaître une fenêtre qui semblait provenir du logiciel antivirus, mais qui vous a semblé fausse?

a) Oui (précisez) : _____

b) Non

c) Je ne suis pas certain (précisez) : _____

11) Pensez-vous que votre portable est actuellement infecté par un virus, un logiciel malveillant ou qu'il présente un problème de sécurité?

a) Oui (précisez) :_____

b) Non

c) Je ne suis pas sûr (précisez) :_____

Les questions 12-20 portent sur le profil de l'utilisateur; c.-à-d., comment l'usager utilise son ordinateur

12) Au cours du dernier mois, quel est en moyenne le nombre d'heures par jour que votre ordinateur est resté allumé et connecté (que ce soit par les connexions filaires ou sans fil) à Internet, en incluant le temps où l'ordinateur n'est pas utilisé?

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

13) Au cours du dernier mois, quel est en moyenne le nombre d'heures par jour que votre ordinateur est resté allumé et non connecté (que ce soit par les connexions filaires ou sans fil) à Internet, en incluant le temps où l'ordinateur n'est pas utilisé?

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

14) Au cours du dernier mois, combien d'heures par jour passez-vous en moyenne sur l'ordinateur à réaliser des tâches sur Internet (naviguer, jeux en ligne, réseaux sociaux, etc.)?

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

15) Au cours du dernier mois, combien d'applications (non utilisées par un navigateur web) ont été installées sur l'ordinateur (par vous ou par quelqu'un d'autre)?

|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 ou plus

16) Comment classifieriez-vous les programmes que vous avez installés sur l'ordinateur au cours du dernier mois?

a) La majorité de ces programmes sont des jeux
b) La majorité de ces programmes ne sont **pas** des jeux
c) Aucun programme n'a été installé au cours du dernier mois
d) Autre (spécifiez) :_________________

17) Comment classifieriez-vous les programmes qui ont été installés par d'autres

personnes que vous sur l'ordinateur au cours du dernier mois?
a) La majorité de ces programmes sont des jeux
b) La majorité de ces programmes ne sont **pas** des jeux
d) Aucune autre personne que moi n'a installé des programmes sur l'ordinateur
e) Je ne sais pas
e) Autre (spécifiez) :_____

18) Au cours du dernier mois, combien de plugins web (logiciel servant à ajouter une fonctionnalité à un navigateur web) ont été installés sur l'ordinateur (par vous ou par quelqu'un d'autre)?

|____|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 ou plus

19) Parmi les applications suivantes, quelle est celle que vous avez utilisée le **plus fréquemment** au cours du dernier mois?
a) Navigateur web (par exemple, Internet Explorer, Firefox, Chrome, etc.)
b) Suite Office (par exemple, MS Word, PowerPoint, Excel, Openoffice, etc)
c) Application courriel non disponible par un navigateur web (par exemple, Outlook, Thunderbird, etc)
d) Jeux
e) Autre (spécifiez): _____

20) Parmi les applications suivantes, quelle est la **deuxième** application que vous avez utilisée le **plus fréquemment** au cours du dernier mois?
a) Navigateur web (par exemple, Internet Explorer, Firefox, Chrome, etc.)
b) Suite Office (par exemple, MS Word, PowerPoint, Excel, Openoffice, etc)
c) Application courriel (par exemple, Outlook, Thunderbird, etc)
d) Jeux
e) Autre (spécifiez): _____

Les questions 21-25 visent à déterminer les habitudes de navigation de l'utilisateur

21) Comment classifieriez-vous les sites web que vous avez visités le plus souvent au cours du dernier mois?
a) Sites de réseaux sociaux ou de forums en ligne (par exemple, Facebook, Myspace, etc)
b) Sites de jeux en ligne
c) Sites pour adultes
d) Sites de vidéos (par exemple, Youtube)
e) Moteurs de recherche
f) Autre (spécifiez): _____

22) Au cours du dernier mois, combien de fois avez-vous utilisé l'ordinateur pour installer des applications en provenance d'Internet?

|____|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 ou plus

23) Au cours du dernier mois, combien de fois avez-vous utilisé l'ordinateur pour télécharger des fichiers vidéo ou audio à partir d'Internet?

|____|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 ou plus

24) Quel(s) navigateur(s) web est(sont) actuellement installé(s) sur votre ordinateur? Sélectionnez tous les choix applicables.

[] Internet Explorer
[] Firefox
[] Chrome
[] Autre (spécifiez): _____

25) Au cours du dernier mois, quel navigateur web avez-vous utilisé le plus souvent?

a) Internet Explorer
b) Firefox
c) Chrome
d) Autre (spécifiez): _____

26) Utilisez-vous actuellement les paramètres par défaut de sécurité et de confidentialité des navigateurs web ou avez-vous apporté des changements? Sélectionnez tous les choix applicables.

[] J'utilise les paramètres par défaut de sécurité et de confidentialité de tous les navigateurs web installés sur le portable

[] J'ai apporté des changements aux paramètres de sécurité et de confidentialité du ou des navigateurs web suivants (précisez) :_____

[] Autre (spécifiez): _____

Les questions 27-28 concernent la variabilité de l'environnement réseau du sujet

27) Au cours du dernier mois, quel est l'endroit principal à partir duquel vous vous êtes connecté à Internet en utilisant votre ordinateur?

a) Domicile
b) Campus universitaire
c) Travail
d) Café
e) Autre (spécifiez): _____

28) Au cours du dernier mois, combien d'endroits différents avez-vous utilisés pour vous connecter à Internet à partir de votre ordinateur?

|____|____|____|____|____|____|____|____|____|____|
1 2 3 4 5 6 7 8 9 10 ou plus

La question 29 tente de déterminer le niveau de participation du sujet aux activités réseaux P2P

29) Au cours du dernier mois, combien de fois avez-vous utilisé les réseaux peer-to-peer (Bittorrent, Limeware, Kazaa, etc.) pour télécharger des fichiers audio, vidéo ou autres sur l'ordinateur?

|____|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 ou plus

Les questions 30-33 visent à déterminer le niveau de vigilance informatique du sujet et la mesure dans laquelle il applique cette dernière afin de sécuriser son ordinateur.

30) Sélectionnez les tâches que vous avez déjà réalisées; si aucune de ces tâches ne s'applique à votre situation, sélectionnez "Aucune de ces réponses".

[] J'ai configuré un pare-feu.
[] J'ai sécurisé un réseau sans-fil.
[] J'ai modifié les paramètres par défaut de sécurité d'un navigateur web.
[] plugins…
[] Aucune de ces réponses

31) Sur une échelle de 1 à 10, sélectionnez la valeur qui représente le mieux votre niveau de préoccupation à l'égard de la sécurité de votre ordinateur. Les valeurs élevées signifient une plus grande préoccupation.

Aucune préoccupation Préoccupation élevée
|____|____|____|____|____|____|____|____|____|____|
1 2 3 4 5 6 7 8 9 10

32) Quel énoncé s'applique le mieux à votre situation?

a) Je ne suis pas préoccupé à l'idée que mon ordinateur puisse être compromis
b) Je suis préoccupé par la sécurité de mon ordinateur mais je ne sais pas quoi faire pour l'empêcher d'être compromis
c) Je sais ce que je dois faire pour sécuriser mon ordinateur mais je suis trop occupé pour agir en conséquence
d) Je sais quoi faire pour sécuriser mon ordinateur et j'agis en conséquence en réalisant les tâches suivantes (spécifiez) :

_____

e) Autre (spécifiez):

_____

33) Quel énoncé parmi les suivants s'applique le mieux à votre situation?

a) Je clique sur les liens et les fichiers joints aux courriels seulement si le contenu m'intéresse

b) Je clique seulement sur les liens et les fichiers joints aux courriels qui proviennent de personnes que je connais

c) Autre (spécifiez) :_____

Les questions 34-35 visent à déterminer les habitudes du sujet à l'égard de ses mots de passe

34) Combien avez-vous approximativement de comptes web protégés par mot de passe?_____

35) Quel énoncé s'applique le mieux à votre situation en ce qui a trait à votre gestion des mots de passe?

a) Je garde seulement 1 à 2 mots de passe pour me connecter à mes comptes

b) J'ai 3 à 5 mots de passe pour me connecter à mes comptes

c) J'ai 6 à 10 mots de passe pour me connecter à mes comptes

d) J'utilise un mot de passe différent pour chaque compte

e) Je n'utilise pas un mot de passe différent pour chaque compte, mais j'utilise plus de 10 mots de passe pour me connecter à mes comptes

f) Autre (spécifiez) : _____

ANNEXE F

QUESTIONNAIRE EN CAS D'INFECTION

Questions pour les sujets dont les portables sont infectés ou soupçonnés d'être infectés

Ces questions apparaîtront sur un formulaire séparé.

Entrez le numéro d'identification du portable: _____

1) Avez-vous observé des comportements étranges dans le fonctionnement de l'ordinateur ?
   a) Oui (spécifiez) :_____
   b) Non
   c) Je ne sais pas

2) Si vous avez répondu oui à question 1, avez-vous une idée de ce qui pourrait être à l'origine de ce problème?
   a) Oui; spécifiez : _____
   b) Non

3) Vous souvenez-vous de la réception d'un message de l'ordinateur ou du logiciel antivirus vous avertissant d'un problème de sécurité?
   a) Oui, le message m'a demandé de réaliser le(s) tâche(s) suivante(s) : _____
   b) Oui, le message m'a seulement informé de la présence d'un problème de sécurité
   c) Non, je ne me souviens pas de la réception d'un message m'avisant d'un problème de sécurité

Si vous avez répondu oui à la question 3, continuez à la question suivante; sinon passez à la question 8.

4) Vous souvenez-vous de ce que vous étiez en train de faire lorsque le message est apparu?
   a) Oui; spécifiez : _____
   b) Non

5) Avez-vous compris ce que le guide-opérateur vous a demandé de faire?
   a) Oui
   b) Non

6) Quelle a été votre réaction?
   a) J'ai suivi la totalité des instructions.
   b) J'ai suivi une partie des instructions.
   c) Je n'ai rien fait.

7) Comment vous êtes-vous senti lorsque vous avez vu le message apparaître? Sélectionnez tous les choix applicables.
   [] J'étais ennuyé par l'interruption
   [] J'étais inquiet concernant la sécurité de l'ordinateur
   [] J'étais confus
   [] J'étais rassuré
   [] Autre (spécifiez) : _____

8) Au cours du dernier mois, avez-vous remarqué une différence dans l'opération du portable?

a) Oui; (spécifiez la nature et si possible une date de début) :_____
b) Non

ANNEXE G

FORMULAIRE DE CONSENTEMENT SUPPLÉMENTAIRE

Formulaire d'information et de consentement supplémentaire

Titre du projet de recherche
Essai clinique de produit antivirus

Préambule
Vous avez été averti que le logiciel antivirus ou que nos outils diagnostiques installés sur votre portable indiquent que ce dernier est probablement infecté. Nous vous demandons par la présente votre autorisation afin de recueillir des données brutes de votre portable afin d'être en mesure d'établir le moyen et la source de l'infection.

Confidentialité
Avec votre consentement, le chercheur responsable ainsi que les membres du personnel recueilleront et consigneront certaines données brutes disponibles sur votre portable. Ces données comprennent :

_____ Les historiques de navigation

_____ Le journal du logiciel *Wireshark* qui contient les identités des systèmes auxquels votre portable s'est connecté

_____ Les fichiers infectés

Tous les renseignements recueillis demeureront strictement confidentiels. Afin de préserver votre identité et la confidentialité des renseignements, vous ne serez identifié que par le numéro de votre portable. Le code reliant votre nom à votre dossier de recherche sera accessible au personnel de recherche impliqué dans le projet tout au long de sa durée. Une fois l'expérience terminée, seul le chercheur responsable aura accès à cette information, et ce pour une durée limitée de 8 mois, après quoi cette information sera supprimée.

Le chercheur responsable et les membres du personnel utiliseront les données brutes à des fins de recherche dans le but de répondre aux objectifs scientifiques du projet décrits dans le formulaire de consentement. Tous les renseignements recueillis au cours du projet de recherche seront conservés dans la zone de haute sécurité du laboratoire SecSI afin d'en assurer la protection. Seuls le chercheur responsable et les membres du personnel auront accès à ces données. Les données brutes de votre dossier seront conservées pendant 1 an par le chercheur responsable. À la fin de cette période, des mesures nécessaires seront prises afin de supprimer adéquatement les données conservées.

Une procédure simple vous sera donnée en fin d'expérience ou en cas de retrait si vous désirez supprimer les informations enregistrées sur votre portable tout au long de l'expérience. Le personnel de recherche sera disponible pour vous guider dans cette procédure, mais vous devrez réaliser vous-même les manipulations.

Les résultats de l'expérience pourront être publiés ou faire l'objet de discussions scientifiques, mais il ne sera pas possible de vous identifier puisque les données sont anonymes.

Identification des personnes-ressources
José M. Fernandez, ing. PhD
Professeur adjoint au Département de Génie Informatique et de Génie Logiciel
Responsable du laboratoire SecSI
Courriel : *jose.fernandez@polymtl.ca*
Tél : (514) 340-4711 p.5433

Carlton Davis, PhD
Stagiaire Postdoctoral au Département de Génie Informatique et de Génie Logiciel
Courriel : *carlton.davis@polymtl.ca*
Tél : (514) 340-5121 p.2063

Fanny Lalonde Lévesque
Étudiante M.Sc.A. au Département de Génie Informatique et de Génie Logiciel
Responsable du soutien technique auprès des sujets recherche
Courriel : *fanny.lalonde-levesque@polymtl.ca*
Tél : (514) 340-4711 p.7172

Pier-Luc St-Onge, M.Sc.A
Analyste au Département de Génie Informatique et de Génie Logiciel
Responsable technique du projet
Courriel : *pier-luc.st-onge@polymtl.ca*
Tél : (514) 340-4711 p.5065

Surveillance des aspects éthiques du projet de recherche
Le comité d'éthique de la recherche (CÉR) de l'École Polytechnique de Montréal a approuvé ce projet de recherche et en assure le suivi. De plus, il approuvera au préalable toute révision et toute modification apportée au formulaire de consentement et au protocole de recherche.

Pour de plus amples informations, vous pouvez contacter directement le président du CÉR de l'École Polytechnique de Montréal :

Bernard Lapierre
Courriel : *bernard.lapierre@polymtl.ca*
Tél : (514) 340-4711 p. 4567

Remerciements
Votre collaboration au projet de recherche est précieuse et nous vous en remercions.

Consentement
J'ai pris connaissance du formulaire de consentement. Je reconnais qu'on m'a expliqué le projet, qu'on a répondu à mes questions et qu'on m'a laissé le temps voulu pour prendre une décision.

Je consens à autoriser la collecte des données brutes aux conditions qui y sont énoncées.

Nom et signature du sujet de recherche Date

J'ai expliqué au sujet de recherche les termes du présent formulaire de consentement et j'ai répondu aux questions qu'il m'a posées.

Nom et signature de la personne qui obtient le consentement Date

ANNEXE H

QUESTIONNAIRE FINAL

Questions pour le sondage à la fin de l'expérience

Ces questions apparaîtront sur un formulaire séparé.

Entrez le numéro d'identification du portable: _____

1) Avez-vous déjà supprimé des informations de l'ordinateur avant de vous présenter aux rencontres mensuelles?
a) Oui
b) Non

2) Combien de fois approximativement avez-vous effacé votre historique de navigation au cours de l'expérience?

|____|____|____|____|____|____|____|____|____|____|____|
0 1 2 3 4 5 6 7 8 9 10 fois ou plus

3) Avez-vous utilisé un mode de navigation privée afin de ne pas générer d'historique?
a) Oui (justifiez) :_____
b) Non

4) Avez-vous modifié votre utilisation du portable sachant vous participiez à une étude?
a) Oui (précisez) :_____
b) Non

5) Est-ce que certaines des réponses que vous avez fournies lors des différents questionnaires ont été influencées par ce que vous pensiez être la bonne réponse et non nécessaire ce qui représentait la situation actuelle?
a) Oui
b) Non

6) Est-ce que certaines questions des sondages précédents auraient influencé votre utilisation du portable?
a) Oui
b) Non

7) Est-ce que certaines des réponses que vous avez fournies lors des questionnaires précédents auraient influencé vos réponses des questionnaires subséquents?
a) Oui
b) Non

8) Est-ce que certaines des réponses que vous avez fournies lors des questionnaires ont été influencées par l'image que pouvait projeter une réponse en particulier?
a) Oui
b) Non

9) Est-ce que votre participation à l'expérience a entraîné des changements quant à votre vigilance en

matière de sécurité informatique?
a) Oui; spécifiez _____
b) Non

10) Avez-vous des préoccupations particulières concernant l'expérience?
a) Oui; spécifiez _____
b) Non

11) Accepteriez-vous de participer à nouveau à une expérience similaire?
a) Oui
b) Non; spécifiez pourquoi _____

12) Commentaires?
_____

ANNEXE I

PROTOCOLE VISITE INITIALE

Protocole pour le déroulement de la rencontre initiale

Ce protocole décrit les étapes à suivre lors de la première rencontre, soit celle où l'ordinateur portable est remis à l'utilisateur.

Les étapes 1 à 7 concernent le consentement du participant
1. Accueillir le participant et lui expliquer le déroulement de la rencontre.
2. Remettre une copie du formulaire d'information et de consentement au participant et lui permettre de lire le formulaire en silence.
3. Répondre aux questions du participant.
4. Préciser au participant :
   - que le paiement de la compensation financière sera effectué par chèque une fois l'expérience complétée.
   - qu'aucune manipulation n'est requise pour l'activation de la garantie auprès de CoopHEC.
5. Déballer l'ordinateur et le présenter au participant.
6. Procéder à la signature des deux copies du formulaire d'information et de consentement.
7. Demander au participant d'inscrire ses informations personnelles sur la feuille de route du participant et sur la feuille de route de l'équipe de recherche.

Les étapes 8 à 10 concernent l'achat de l'ordinateur
8. S'assurer que le participant a bien le paiement de 350$.
9. Amener une des feuilles de route et se présenter avec le participant :
   - au bureau de Chantal Faubert si la rencontre est entre 8h et 9h du matin.
   - au bureau de Chantal Balthazard si la rencontre est après 9h
   - au bureau de Élisabeth Bernier si Chantal Balthazard n'est pas disponible.
10. Une fois le paiement effectué, retourner au local avec le participant.

11. Inviter le participant à remplir le questionnaire initial en ligne.
12. Pendant que le participant répond au questionnaire, prendre en note dans le cahier associé les éléments suivants :
   - Numéro de série de l'ordinateur portable
   - Adresse MAC locale de l'ordinateur portable
   - Adresse MAC sans-fil de l'ordinateur portable
13. Ouvrir le Planificateur de tâches Windows et activer les tâches suivantes :
   - ClinicalTrial_History
   - ClinicalTrial_TShark
   - ClinicalTrial_Applications
   - ClinicalTrial_Locations
   - ClinicalTrial_Uptime
   - ClinicalTrial_Connection
14. Changer le nom de l'ordinateur portable pour « USERXX » où XX est le numéro d'identification du participant. Par exemple, pour le participant 6, mettre « USER06 » comme nom d'ordinateur. Mentionner au participant que le nom de son ordinateur a été changé.
15. Redémarrer l'ordinateur et le connecter à Internet en local.
16. Ouvrir la console de Officescan et aller dans « Networked Computers », sélectionner « Connection Verification » et cliquer sur « Verify Now ».
17. Toujours dans la console de Officescan, rester dans « Networked Computers » et sélectionner « Client Management ». Cliquer sur « refresh » jusqu'à l'apparition de l'ordinateur « USERXX » dans la liste des clients.
18. Valider les réponses du participant au questionnaire.
19. Signer la feuille de route du participant et lui demander de signer la feuille de route de l'équipe de recherche. Préciser au participant qu'il est nécessaire d'amener la feuille de route ainsi que l'ordinateur lors des rencontres mensuelles.
20. Remercier le participant de sa collaboration.

ANNEXE J

PROTOCOLE VISITE MENSUELLE

Protocole pour le déroulement des rencontres mensuelles

Ce protocole décrit les étapes à suivre lors des rencontres mensuelles.

1. Accueillir le participant et lui expliquer le déroulement de la rencontre.
2. Inviter le participant à remplir le questionnaire mensuel en ligne et valider les réponses lorsqu'il aura terminé.
3. Récupérer l'ordinateur, le brancher et l'ouvrir.
4. Connecter l'ordinateur à internet par câble réseau.
5. Ajouter l'ordinateur à la liste des clients sur le serveur s'il n'est pas déjà visible.
6. Modifier les options de dossiers et de recherche afin d'afficher les dossiers et les fichiers cachés.

Les étapes 7 à 8 concernent la vérification du fonctionnement des scripts automatiques

7. Ouvrir le dossier « C:/ClinicalTrial » et vérifier que tous les fichiers et dossiers sont présents.
8. Ouvrir le planificateur de tâches Windows et vérifier que toutes les tâches sont actives et fonctionnelles.

Les étapes 9 à 11 concernent l'exécution de Sigcheck

9. Dans le dossier « C:\ClinicalTrial\scans\Sigcheck-data», créer un dossier pour le mois analysé (et non le mois de la rencontre).
10. Lancer l'exécution de Sigcheck avec la commande suivante :
`« sigcheck -e -i -h -r -s C :\ > C:\ClinicalTrial\scans\Sigcheck-data\MM-YYYY\Sigcheck-JJ-MM-AAAA.txt»`
11. Une fois l'exécution de la commande terminée, lancer le script « sigcheck-output-analysis.pl» et suivre les instructions.

Les étapes 12 à 39 concernent l'exécution des scripts de compilation

12. Insérer dans le lecteur CD/DVD le DVD contenant les scripts.

13. Lancer l'exécution du script « url_classification.pl » et suivre les instructions.
14. Vérifier que l'exécution s'est bien déroulée et qu'un fichier « Navigateur-classification-info-MM-AAAA.txt » a été créé pour chaque navigateur installé sur l'ordinateur dans le dossier « Navigateur/MM-AAAA » associé.

15. Lancer l'exécution du script « uptime_connection_month.pl » et suivre les instructions.
16.Vérifier que l'exécution s'est bien déroulée et qu'un fichier «connectivity-info-MM-AAAA.txt » a bien été créé dans le dossier « C:/ClinicalTrial/data/connection-time/MM-AAAA/ ».

17. Lancer l'exécution du script « uptime_month.pl » et suivre les instructions.
18.Vérifier que l'exécution s'est bien déroulée et qu'un fichier « uptime-MM-AAAA.txt » a bien été créé dans le dossier « C:/ClinicalTrial/data/uptime/MM-AAAA/ ».

19. Créer le dossier « C:/ClinicalTrial/data/updates-info/MM-AAAA/ » s'il n'existe pas.
20. Exécuter l'application « UpdateChecker ». Une fois le log ouvert dans un navigateur web, enregistrer la page sous « updateChecker-output-DD-MM-AAAA.htm» avec l'option (seulement le contenu html) dans le dossier associé « C:/ClinicalTrial/data/updates-info/MM-AAAA/ » du mois qui doit être analysé et non du mois courant.
21. Lancer l'exécution du script « updates.pl » et suivre les instructions.
22. Vérifier que l'exécution s'est bien déroulée et que les fichiers « updates-info-JJ-MM-AAAA,txt » ont bien été créés dans le dossier associé.

23. Lancer l'exécution du script « tshark_sort.pl » et suivre les instructions.
24. Vérifier l'exécution s'est bien déroulée en regardant les fichiers dans le dossier « C:/ClinicalTrial/data/logs »

25. Lancer l'exécution du script « nb_host.pl » et suivre les instructions.
26. Vérifier que l'exécution s'est bien déroulée et que le fichier « num-host-MM-AAAA.txt » a bien été dans le dossier « C:/ClinicalTrial/logs/MM-AAAA » associé.

29. Ouvrir le navigateur Chrome s'il est installé sur l'ordinateur et écrire « about :plugins » dans la barre de navigation.
30. Sauvegarder uniquement le contenu html de la page sous « Chrome-plugins-MM-AAAA.htm » dans le dossier « C:/ClinicalTrial/data/Chrome/MM-AAAA ».
31. Ouvrir le navigateur Firefox s'il est installé sur l'ordinateur et écrire « about :plugins » dans la barre de navigation.
32. Copier l'ensemble du contenu et le coller dans un document texte. Sauvegarder sous « Firefox-plugins-MM-AAAA.txt » dans le dossier « C:/ClinicalTrial/data/Firefox/MM-AAAA ».
33. Ouvrir internet explorer et sélectionner « Gérer les modules complémentaires » dans le menu « Outils ».
34. Modifier l'affichage pour voir tous les modules.
35. Sélectionner le contenu et coller dans un document texte. Sauvegarder sous « IE-plugins-MM-AAAA.txt » dans le dossier « C:/ClinicalTrial/data/IE/MM-AAAA/ ».

36. Lancer l'exécution du script « browser_frequency.pl » et suivre les instructions.
37. Vérifier que l'exécution s'est bien déroulée et qu'un fichier « Navigateur-info-MM-AAAA.txt » a été créé pour chaque navigateur installé sur l'ordinateur dans le dossier « Navigateur/MM-AAAA » associé.

38. Lancer l'exécution du script « type_files.pl » et suivre les instructions.
39. Vérifier que l'exécution s'est bien déroulée et qu'un fichier « Navigateur-download-info-MM-AAAA.txt » a été créé pour chaque navigateur installé sur l'ordinateur dans le dossier « Navigateur/MM-AAAA » associé.

Les étapes 40 à 43 concernent SpyBHORemover

40. Dans le dossier « C:/ClinicalTrial/scans/Spybho-data/» créer le dossier associé au mois à analyser.
41. Lancer l'exécution de SpyBHORemover et enregistrer le journal sous « Spybho-JJ-MM-AAAA.html » dans le dossier « C:/ClinicalTrial/scans/Spybho-data/MM-AAAA» approprié.
42. Si un des BHO est identifié comme étant « Dangereux », l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer par la liste O2/O3 du site *www.systemlookup.com* ou par toutes autres ressources jugées nécessaires que le BHO est bien associé à un malware.
43. Si un des BHO est identifié comme étant « Suspect » ou « Besoin d'analyse », noter le CLSID du BHO en question et vérifier si ce dernier est associé à un malware par la liste O2/O3 du site *www.systemlookup.com* ou par toutes autres ressources jugées nécessaires. Si le BHO se révèle comme étant associé à un malware, l'ordinateur peut être considéré comme étant infecté.

Les étapes 44 à 48 concernent SpyDLLRemover

44. Dans le dossier « C:/ClinicalTrial/scans/Spydll-data/» créer le dossier associé au mois à analyser.
45. Ouvrir l'application SpyDLLRemover et changer les paramètres pour afficher les menaces dangereuses, suspectes et qui ont besoin d'analyse.
46. Lancer le scan et enregistrer le journal sous « Spydll-JJ-MM-AAAA.html » dans le dossier « C:/ClinicalTrial/scans/Spydll-data/MM-AAAA» approprié.
47. Si un fichier est considéré comme étant « Dangereux », l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer par le site *www.processlibrary.com* ou par toutes autres ressources jugées nécessaires que le fichier est bien associé à un malware.
48. Si un des fichiers est identifié comme étant « Suspect » ou « Besoin d'analyse », vérifier si ce dernier est associé à un malware par le site *www.processlibrary.com* ou par toutes autres ressources jugées nécessaires. Si le fichier se révèle comme étant associé à un malware, l'ordinateur peut être considéré comme étant infecté.

Les étapes 49 à 55 concernent Autoruns

49. Dans le dossier « C:/ClinicalTrial/scans/Autoruns-data/» créer le dossier associé au mois à analyser.
50. Ouvrir une fenêtre de console à partir du dossier « C:\ClinicalTrial\bin\Autoruns » et lancer la commande suivante :
`« autorunsc –a –v -x > C:\ClinicalTrial\scans\Autoruns-data\MM-AAAA\Autoruns-JJ-MM-AAAA.xml»`
51. Lancer l'exécution du script « autoruns.pl » et suivre les instructions.
52. Vérifier que l'exécution du script s'est bien déroulée.
53. Dans le dossier « C:\ClinicalTrial\scans\Autoruns-data\MM-AAAA» voir si un fichier « Autoruns-Dangerous.txt » a été créé. Dans un tel cas, l'ordinateur peut être considéré comme étant potentiellement infecté. Vérifier sur les sites *www.pacs-portal.co.uk/startup_search.php*, *www.processlibrary.com* et *www.systemlookup.com* pour confirmer s'il y a bien infection.
54. Dans le dossier « C:\ClinicalTrial\scans\Autoruns-data\MM-AAAA» voir si un fichier « Autoruns-Suspicious.txt » a été créé. Advenant que oui, vérifier sur les sites *www.pacs-portal.co.uk/startup_search.php*, *www.processlibrary.com* et *www.systemlookup.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.
55. Dans le dossier « C:\ClinicalTrial\scans\Autoruns-data\MM-AAAA» voir si un fichier « Autoruns-Unrateds.txt » a été créé. Advenant que oui, vérifier sur les sites *www.pacs-portal.co.uk/startup_search.php*, *www.processlibrary.com* et *www.systemloopup.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.

Les étapes 56 à 63 concernent Process Explorer

56. Dans le dossier « C:/ClinicalTrial/scans/ProcessExplorer-data/» créer le dossier associé au mois à analyser.
57. Ouvrir l'application ProcessExplorer. Modifier l'affichage pour obtenir exactement les colonnes suivantes : Process, PID, Description, Company Name, Path, Verified Signer.
58. Sauvegarder le journal sous « C:/ClinicalTrial/scans/ProcessExplorer-data/MM-AAAA/ProcessExplorer-JJ-MM-AAAA.txt».
59. Lancer l'exécution du script « process_explorer.pl » et suivre les instructions.
60. Vérifier que l'exécution s'est bien déroulée en regardant dans le dossier « C:/ClinicalTrial/scans/ProcessExplorer-data/MM-AAAA/» que les différents fichiers ont été créés.
61. Dans le dossier « C:\ClinicalTrial\scans\ProcessExplorer-data\MM-AAAA» voir si un fichier « ProcessExplorer-Dangerous.txt » a été créé. Dans un tel cas, l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer par le site *www.processlibrary.com* ou par toutes autres ressources jugées nécessaires que le fichier est bien associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.
62. Dans le dossier « C:\ClinicalTrial\scans\ProcessExplorer-data\MM-AAAA» voir si un fichier « ProcessExplorer-Suspicious.txt » a été créé. Advenant que oui, vérifier manuellement sur le site *www.processlibrary.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.
63. Dans le dossier « C:\ClinicalTrial\scans\ProcessExplorer-data\MM-AAAA» voir si un fichier « ProcessExplorer-Unrated.txt » a été créé. Advenant que oui, vérifier manuellement sur le site *www.processlibrary.com* ou par toutes autres ressources jugées nécessaires si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.

Les étapes 64 à 88 concernent Hijackthis

64. Dans le dossier « C:/ClinicalTrial/scans/Hijackthis-data/» créer le dossier associé au mois à analyser.
65. Ouvrir l'application Hijackthis en tant qu'administrateur et lancer un scan.

66. Sauvegarde le journal dans le fichier « C:/ClinicalTrial/scans/Hijackthis-data/MM-AAAA/Hijackthis-JJ-MM-AAAA.log».
67. Lancer l'exécution du script « hijackthis.pl » et suivre les instructions.
68. Vérifier que l'exécution s'est bien déroulée en vérifiant si les fichiers associés ont été créés dans le dossier « C:/ClinicalTrial/scans/Hijackthis-data/MM-AAAA».
69. Si un fichier « Hijackthis-F0.txt » a été créé, l'ordinateur est possiblement infecté. Dans un tel cas, analyser manuellement le fichier avec le site *www.processlibrary.com* et si nécessaire *www.virustotal.com*.
70. Si un fichier « Hijackthis-F1.txt » a été créé, l'ordinateur est possiblement infecté. Dans un tel cas, analyser manuellement le fichier avec le site *www.processlibrary.com* et si nécessaire avec *www.virustotal.com*.
71. Si un fichier « Hijackthis-F2.txt » a été créé, analyser manuellement le contenu du fichier avec le site *www.processlibrary.com* et si nécessaire avec *www.virustotal.com* ou toutes autres ressources jugées nécessaires pour vérifier si le fichier est associé à un malware.
72. Si un fichier « Hijackthis-F3.txt » a été créé, analyser manuellement le fichier avec le site *www.processlibrary.com* et si nécessaire avec *www.virustotal.com* ou toutes autres ressources jugées nécessaires pour vérifier si le fichier est associé à un malware.
73. Si un fichier « Hijackthis-O1.txt » a été créé, l'ordinateur est peut-être infecté. Analyser manuellement l'entrée à l'aide du site *www.malwaredomainlist.*com pour voir si l'ordinateur est infecté.
74. Si un fichier « Hijackthis-O2.txt » a été créé, analyser manuellement chacun des BHO par le CLSID à l'aide du site *www.systemlookup.com* pour voir si un BHO est associé avec un malware. Dans un tel cas, l'ordinateur est infecté.
75. Si un fichier « Hijackthis-O3.txt » a été créé, analyser manuellement chacune des barres d'outils par le CLSID à l'aide du site *www.systemlookup.com* pour voir si une barre d'outils est associé avec un malware. Dans un tel cas, l'ordinateur est infecté.
76. Si un fichier « Hijackthis-O4-Dangerous.txt » a été créé, l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer à l'aide des sites *www.processlibrary.com*, *www.pacs-portal.co.uk/startup_search.php* et *www.systemlookup.com* si le fichier est associé avec un malware.
77. Si un fichier « Hijackthis-O4-Suspicious.txt » a été créé, vérifier manuellement sur les sites *www.processlibrary.com*, *www.pacs-portal.co.uk/startup_search.php* et *www.systemlookup.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.
78. Si un fichier « Hijackthis-O10.txt » a été créé, l'ordinateur est possiblement infecté. Dans un tel cas, analyser manuellement le fichier avec le site *www.systemlookup.com* pour confirmer s'il y a infection.
79. Si un fichier « Hijackthis-URL-Dangerous.txt » a été créé, l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer que l'adresse url est bien malveillante par le site *www.malwaredomainlist.com*.
80. Si un fichier « Hijackthis-O17.txt » a été créé, l'ordinateur est possiblement infecté. Dans un tel cas, analyser manuellement l'entrée avec le site *www.malwaredomainlist.com* pour confirmer si une adresse IP ou un site web est associé à un site malveillant.
81. Si un fichier « Hijackthis-O18.txt » a été créé, analyser manuellement le contenu du fichier avec le site *www.systemlookup.com* pour déterminer si un protocole est associé à un malware.
82. Si un fichier « Hijackthis-O19.txt » a été créé, l'ordinateur est possiblement infecté. Dans un tel cas, analyser manuellement le contenu du fichier avec le site *www.virustotalcom* pour confirmer si le fichier .css est malveillant.
83. Si un fichier « Hijackthis-O20.txt » a été créé, analyser manuellement le contenu du fichier avec le site *www.systemlookup.com* et avec le site *www.processlibrary.com* pour déterminer si un fichier est associé à un malware.
84. Si un fichier « Hijackthis-O21.txt » a été créé, analyser manuellement le contenu du fichier à l'aide du CLSID associé par le site *www.systemlookup.com* et avec le site *www.processlibrary.com* si nécessaire pour déterminer si un fichier est associé à un malware.
85. Si un fichier « Hijackthis-O22.txt » a été créé, analyser manuellement le contenu du fichier à l'aide du CLSID associé par le site *www.systemlookup.com* et avec le site *www.processlibrary.com* si nécessaire pour déterminer si un fichier est associé à un malware.
86. Si un fichier « Hijackthis-O23-Dangerous.txt » a été créé, l'ordinateur peut être considéré comme étant potentiellement infecté. Confirmer par les sites *www.processlibrary.com*, *www.pacs-portal.co.uk/startup_search.php* et *www.systemlookup.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.

87. Si un fichier « Hijackthis-O23-Suspicious.txt » a été créé, vérifier manuellement sur les sites *www.processlibrary.com*, *www.pacs-portal.co.uk/startup_search.php* et *www.systemlookup.com* si le fichier est associé à un malware. Si oui, l'ordinateur peut être considéré comme étant infecté.
88. Si un fichier « Hijackthis-O24.txt » a été créé, analyser manuellement le contenu du fichier avec le site *www.virustotal.com* pour déterminer si un fichier est associé à un malware.

Les étapes 89 à 91 concernent WinPrefetchView

89. Dans le dossier « C:/ClinicalTrial/scans/WinPrefetchView-data/» créer le dossier associé au mois à analyser.
90. Ouvrir une fenêtre de console en **mode administrateur** à partir de « C:\ClinicalTrial\bin\WinPrefetchView» et lancer la commande suivante :
« `winprefetchview /stext «C:\ClinicalTrial\scans\WinPrefetchView-data\MM-AAAA\WinPrefetchView-data-JJ-MM-AAAA.txt»` »
91. Ouvrir une fenêtre de console en **mode administrateur** à partir de « C:\ClinicalTrial\bin\WinPrefetchView» et lancer la commande suivante :
« `winprefetchview /prefetchfile C:\Windows\Prefetch\NTOSBOOT-b00DFAAD.pf/stext «C:\ClinicalTrial\scans\WinPrefetchView-data\MM-AAAA\WinPrefetchView-bootinfo-data-JJ-MM-AAAA.txt»` »

Les étapes 92 à 93 concernent Whatchanged

92. Dans le dossier « C:/ClinicalTrial/scans/WhatChanged-data/» créer le dossier associé au mois à analyser.
93. Exécuter WhatChanged et sélectionner l'option "Scan Registry" ainsi que les quatre sous-options (CLASSES ROOT, LOCAL MACHINE, CURRENT USER, USERS). Cliquer sur « STEP #1 SNAPSHOT » et attendre la fin de l'exécution. Une fois l'analyse complétée, copier les fichiers Whatchanged_Snapshot1_Registry_HKCR, Whatchanged_Snapshot1_Registry_HKCU, Whatchanged_Snapshot1_Registry_HKLM, Whatchanged_Snapshot1_Registry_HKU dans C:\ClinicalTrial\scans\Whatchanged-data\Whatchanged_HKCR_DD-MM-YYYY.txt, C:\ClinicalTrial\scans\WhatChanged-data\ Whatchanged_HKCU_DD-MM-YYYY.txt, C:\ClinicalTrial\scans\WhatChanged-data\scans\Whatchanged_HKLM_DD-MM-YYYY.txt, C:\ClinicalTrial\scans\WhatChanged-data\scans\Whatchanged_HKU_DD-MM-YYYY.txt

Les étapes 94 à 95 concernent Sigcheck

94. Vérifier que l'exécution du script s'est bien déroulée.
95. Si un fichier « Sigcheck-Mismatch-Signature.txt» a été créé, l'ordinateur est peut-être infecté. Vérifier par toutes ressources jugées nécessaires si le fichier concerné est associé à un malware ou non. Si oui, l'ordinateur peut être considéré comme étant infecté.

Les étapes 96 à 116 concernent la copie des données statistiques

96. Créer un dossier « IDXX-MM-AAAA » (où XX est le numéro du portable) sur le disque C:/ et y créer les sous dossiers « C:/IDXX-MM-AAAA/data », « C:/IDXX-MM-AAAA/logs » et « C:/IDXX-MM-AAAA/scans ».
97. Créer un sous dossier « C:/IDXX-MM-AAAA/data /application-info » et y copier le dossier « C:/ClinicalTrial/data/application-info/MM-AAAA ».
98. Créer un sous dossier « C:/IDXX-MM-AAAA/data /connection-time » et y copier le dossier « C:/ClinicalTrial/data/connection-time/MM-AAAA ».
99. Créer un sous dossier « C:/IDXX-MM-AAAA/data /updates-info » et y copier le dossier « C:/ClinicalTrial/data/updates-info/MM-AAAA ».

100. Créer un sous dossier « C:/IDXX-MM-AAAA/data /uptime » et y copier le dossier « C:/ClinicalTrial/data/uptime/MM-AAAA » ainsi que le fichier « C:/ClinicalTrial/data/uptime/uptime.txt »
101. Créer un sous dossier «C:/IDXX-MM-AAAA/data / connection-info » et y copier le dossier « C:/ClinicalTrial/data/connection-info/MM-AAAA ».
102. En fonction des navigateurs installés, créer un sous dossier « C:/IDXX-MM-AAAA/data /Navigateur » et y créer un autre sous dossier « MM-AAAA ». Copier les fichiers suivants : « Navigateur-classification-MM-AAAA.txt», « Navigateur-download-MM-AAAA.txt », « Navigateur-info-MM-AAAA.txt », « Navigateur-plugins-MM-AAAA.txt ».
103. Créer un sous dossier « C:/IDXX-MM-AAAA/logs /MM-AAAA » et y copier le fichier « C:/ClinicalTrial/logs/ MM-AAAA/num-host-11-2011.txt ».
104. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /Autoruns-data » et y copier le dossier « C:/ClinicalTrial/scans/Autoruns-data/MM-AAAA/».
105. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /Hijackthis-data » et y copier le dossier « C:/ClinicalTrial/scans/Hijackthis-data/MM-AAAA/».
106. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /ProcessExplorer-data » et y copier le dossier « C:/ClinicalTrial/scans/ProcessExplorer-data/MM-AAAA/».
107. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /Spybho-data » et y copier le dossier « C:/ClinicalTrial/scans/Spybho-data/MM-AAAA/».
108. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /Spydll-data » et y copier le dossier « C:/ClinicalTrial/scans/Spydll-data/MM-AAAA/».
109. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /Sigcheck-data » et y copier le dossier « C:/ClinicalTrial/scans/Sigcheck-data/MM-AAAA/».
110. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /WinPrefetchView-data » et y copier le dossier « C:/ClinicalTrial/scans/WinPrefetchView-data/MM-AAAA/».
111. Créer un sous dossier « C:/IDXX-MM-AAAA/scans /WhatChanged-data » et y copier le dossier « C:/ClinicalTrial/scans/WhatChanged-data/MM-AAAA/».

112. Graver l'ensemble du dossier « C:/IDXX-MM-AAAA » sur un DVD vierge et identifié ce dernier avec : « ID :XX MM-AAAA » ainsi qu'avec « Données statistiques ».
113. Supprimer le dossier « C:/IDXX-MM-AAAA » et supprimer à nouveau le dossier dans la corbeille.

Suivre les instructions du protocole en cas d'infection si l'ordinateur est infecté ou soupçonné de l'être.

114. Modifier les options de dossiers et de recherche afin de ne pas afficher les dossiers et les fichiers cachés.
115. Signer la feuille de route du participant et lui demander de signer la feuille de route de l'équipe de recherche.
116. Remercier le participant de sa collaboration.

ANNEXE K

PROTOCOLE EN CAS D'INFECTION

Protocole pour le déroulement des rencontres mensuelles en cas d'infection

Ce protocole décrit les étapes à suivre lors des rencontres mensuelles si l'ordinateur est infecté ou soupçonné d'être infecté.

1. Expliquer au participant que l'ordinateur est infecté ou possiblement infecté.
2. Présenter au participant le formulaire d'information et de consentement supplémentaire et répondre à ces questions.
3. Poursuivre si le participant donne son consentement. Sinon, lui conseiller de lancer un scan de l'antivirus pour désinfecter son ordinateur.

Les étapes 4 à 6 concernent la complétion de la fiche d'infection

4. Insérer dans le lecteur CD/DVD le DVD contenant la copie de la fiche d'infection et ouvrir le fichier.
5. Créer un dossier « IDXX-MM-AAAA » (où XX est le numéro du portable) sur le disque C:/ et y enregistrer la fiche sous le nom « InfectionNXX-MM-AAAA.txt » (où XX est le numéro de l'infection).
6. Remplir les différentes sections de la fiche. Dans le cas d'une clef de registre, la dernière date de modification peut être obtenue par l'outils RegScanner. Simplement installer l'application à partir du DVD prévu à cet effet.

Les étapes 7 à 8 concernent la compilation des données supplémentaires

7. Lancer l'exécution du script « url_month.pl » et suivre les instructions.
8. Vérifier que l'exécution du script s'est bien déroulée en regardant si le fichier « Navigateur-history-MM-AAAA.txt » a bien été créé dans le dossier « C:/ClinicalTrial/data/Navigateur/MM-AAAA »

Les étapes 9 à 13 concernent la copie des données supplémentaires

9. Créer un dossier « IDXX-MM-AAAA » (où XX est le numéro du portable) sur le disque C:/ et y créer les sous dossiers « C:/IDXX-MM-AAAA/data », « C:/IDXX-MM-AAAA/logs » et « C:/IDXX-MM-AAAA/scans ».
10. Créer un sous dossier « C:/IDXX-MM-AAAA/data /Navigateur » pour chaque navigateur installé et y copier le fichier « C:/ClinicalTrial/data/Navigateur/MM-AAAA/Navigateur-history-MM-AAAA.txt ».
11. Créer un sous dossier « C:/IDXX-MM-AAAA/logs /MM-AAAA » et y copier tous les fichiers « output-tshark-JJ-MM-AAAA.txt ».

12. Graver l'ensemble du dossier « C:/IDXX-MM-AAAA » sur un DVD vierge et identifier ce dernier avec : « ID :XX MM-AAAA » ainsi qu'avec la mention« Données brutes ».
13. Supprimer les sous dossiers « C:/IDXX-MM-AAAA/data », « C:/IDXX-MM-AAAA/logs » et « C:/IDXX-MM-AAAA/scans » et les supprimer à nouveau dans la corbeille.

Les étapes 14 à 21 concernent le chiffrement de fichier(s) malicieux

14. Installer Gpg4win sur l'ordinateur à l'aide du DVD prévu à cet effet.
15. Lors de l'installation, il est important de cocher seulement « Gpg4win » et « Kleopatra ».
16. Une fois l'installation terminée, retirer le DVD du lecteur et insérer le DVD qui contient la clef.
17. Ouvrir Kleopatra et cliquer sur « Import certificate ».
18. Choisir la clef désignée sur le DVD.
19. Une fois la clef importée, cliquer sur « Files » et choisir « Sign/Encrypt files… »
20. Sélectionner le fichier désiré et laisser les options par défaut.
21. Vérifier si l'encryption a bien fonctionné.

Les étapes 22 à 26 concernent la copie de(s) fichier(s) malicieux

22. Créer le sous dossier « C:/IDXX-MM-AAAA/malware ».
23. Copier le ou les fichiers chiffré(s) dans le dossier « C:/IDXX-MM-AAAA/malware ».
24. Copier l'ensemble des fichiers « InfectionNXX-MM-AAAA.txt » dans le dossier « C:/IDXX-MM-AAAA/ ».
25. Graver l'ensemble du dossier « C:/IDXX-MM-AAAA » sur un DVD vierge et identifié ce dernier avec : « ID :XX MM-AAAA » ainsi qu'avec « Malware ».
26. Supprimer le dossier « C:/IDXX-MM-AAAA » et supprimer à nouveau le dossier dans la corbeille.