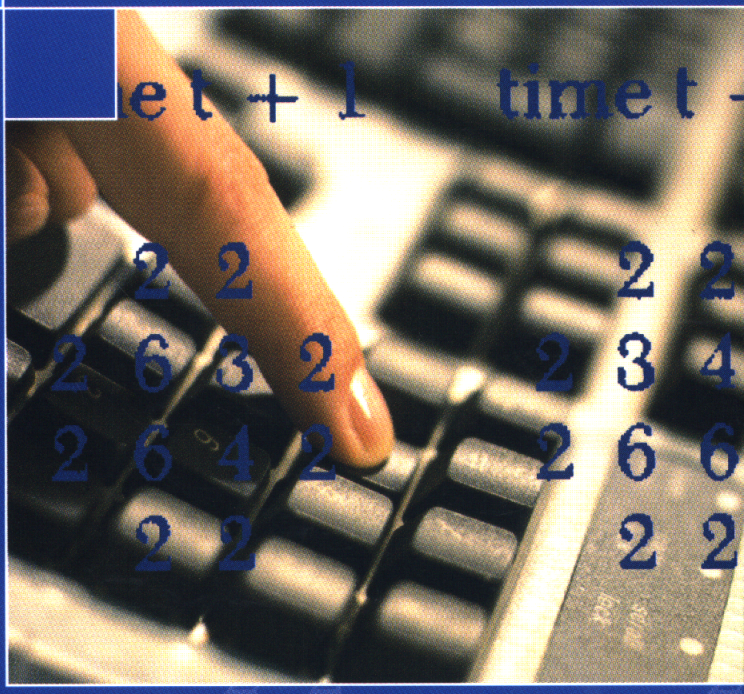
avec CD-ROM

Éric Filiol

Collection IRIS
dirigée par Nicolas Puech



Les virus informatiques : théorie, pratique et applications

BIBLIOTHEQUE DU CERIST

Springer

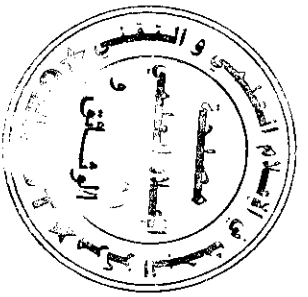
INRIA

Les virus informatiques : théorie, pratique et applications

Springer
Paris
Berlin
Heidelberg
New York
Hong Kong
Londres
Milan
Tokyo

Éric Filiol

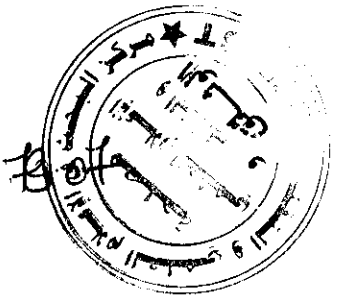
Les virus informatiques : théorie, pratique et applications



Springer

Éric Filiol
Chef du laboratoire de virologie et cryptologie
École Supérieure et d'Application des Transmissions
B.P. 18
35998 Rennes Armées

et INRIA-Projet Codes



ISBN : 2-287-20297-8

Imprimé en France
Springer-Verlag France est membre du groupe Springer Science + Business Media

SPIN: 10966136

Maquette de couverture : Jean-François MONTMARCHÉ

À ma femme Laurence,
à mon fils Pierre,
à mes parents,
à Fred Cohen,
à Mark Allen Ludwig

Collection IRIS

Direction éditoriale : Nicolas PUECH

Maître de conférences

Département Informatique et Réseaux

École Nationale Supérieure des Télécommunications

46, rue Barrault – 75013 Paris

Ouvrages déjà parus :

— A. Quateroni, R. Sacco, F. Saleri. *Méthodes numériques pour le calcul scientifique. Programmes en MATLAB. 2000.*

— J.P. Chancelier, F. Delebecque, C. Gomez, M. Goursat, R. Nikoukhah, S. Steer. *Introduction à SCILAB. 2002.*

— F. Maltey. *Calcul formel avec MUPAD. 2002.*

— A. Meier. *Introduction pratique aux bases de données relationnelles. 2002.*

— B. Goossens. *Architecture et micro-architecture des processeurs. 2002.*

Préface

« Virus don't harm, ignorance do »
herm1t

« ... I am convinced that computer viruses are not evil and that programmers have a right to create them, possess them and experiment with them ... truth seekers and wise men have been persecuted by powerful idiots in every age ... »
Mark A. Ludwig

Tout individu a droit à la liberté d'opinion et d'expression, ce qui implique le droit de ne pas être inquiété pour ses opinions et celui de chercher, de recevoir et de répandre, sans considérations de frontières, les informations et les idées par quelque moyen d'expression que ce soit.
Article 19 - Déclaration universelle des droits de l'Homme

Cet ouvrage traite des « *virus informatiques* », d'un point de vue théorique mais également d'un point de vue pratique et technique le code source de plusieurs virus y est détaillé, décortiqué et commenté. Les « *applications* » utilisant de tels programmes malicieux sont également présentées, cet aspect n'étant quasiment jamais considéré dans les rares ouvrages consacrés aux virus.

Pourquoi un tel livre qui pourrait sembler à certains provocant ? La provocation n'est assurément pas le but. Depuis une petite décennie, force est de constater combien la lutte antivirale connaît de plus en plus de difficultés à s'organiser et à réagir, notamment face aux attaques virales de ces trois ou quatre dernières années. Les vers récents, *Sapphire*, *Blaster* et *Sobig-F* illustrent parfaitement cette situation. Ces attaques, aux effets souvent planétaires, paraissent, pour les utilisateurs qui y sont confrontés mais également aux yeux du grand public, prendre de cours, chaque fois, le monde des éditeurs d'antivirus. Le besoin de faire sortir la lutte antivirale du cadre quasi-confidentiel dans laquelle elle est confinée, se fait sentir de plus en plus. La complexité des problèmes liés à la lutte contre les virus, et en même temps la rareté des ouvrages techniques consacrés à la virologie informatique science qui évolue en permanence militent en faveur d'un tel ouvrage.

En fait, ce livre s'adresse essentiellement aux professionnels de l'informatique ou, au minimum, aux passionnés de cette science ou technique, qui souhaitent acquérir une vision claire et indépendante de ce que sont les virus, et des risques, mais aussi des possibilités, qu'ils représentent. Il ne s'adresse en aucun cas aux « acteurs contestables » de l'informatique que la presse généraliste, écrite ou audio-visuelle, tend à idéaliser et à parer d'un savoir mystérieux, autant que génial. Ces pirates ou autres malfaisants informatiques n'ont de seule motivation que de nuire et leurs exactions, si elles sont immatérielles dans les moyens, sont malheureusement bien réelles, en termes de préjudices. Il était donc nécessaire d'apporter quelques clefs de la connaissance dans le domaine de la virologie informatique, de montrer combien il est faux et dangereux de considérer ces pirates comme des « génies » de l'informatique.

À de rares exceptions près, la grande majorité d'entre eux se contente d'utiliser les créations des autres et ne possède, finalement, que de piètres connaissances dans le domaine. Leur bêtise ne fait que contribuer à jeter l'opprobre sur un domaine de connaissance passionnant. Le respect d'autrui passe par le savoir. Science sans conscience n'est qu'ignorance et ruine de l'âme.

Le problème actuel vient du fait que les utilisateurs (dans son acception la plus large, ce qui inclut les administrateurs) sont condamnés d'une part à faire confiance aux éditeurs d'antivirus et à leurs produits, et d'autre part, à subir, presque sans espoir, les virus programmés par d'autres. L'informatique devait normalement libérer l'Homme. La réalité est toute autre. Il n'est pas concevable que le savoir informatique (les virus en l'espèce qui nous occupe)

soit la chasse gardée de quelques professionnels, dans un but commercial, au détriment de ceux qui ne le possèdent pas.

Le but de ce livre est donc d'initier les utilisateurs aux virus afin qu'ils comprennent les techniques de base mises en œuvre par ces programmes particuliers. En fait, la virologie informatique n'est qu'une branche de l'intelligence artificielle, elle-même partie à la fois des mathématiques et de la science informatique. Les virus ne sont que de simples programmes, aux propriétés certes particulières. Trop longtemps marqués du sceau de l'infamie, ils sont pourtant une réalité qui s'impose à nous avec force mais plus encore, leur intérêt, pour d'éventuelles applications, a été systématiquement et volontairement passé sous silence.

Que cela fasse plaisir ou non, que cela contrarie ou non certains intérêts, les virus sont appelés à jouer un rôle important dans un avenir proche. Le but est donc d'initier suffisamment les utilisateurs pour qu'ils acquièrent une certaine autonomie, notamment dans le domaine de la lutte antivirale, et de pouvoir agir même quand les antivirus échouent. L'enseignement de la virologie informatique commence à timidement s'organiser. L'université de Calgary, au Canada, offre depuis l'automne 2003 un cours sur le sujet, non sans provoquer une vive réaction, de certains acteurs de la communauté antivirale (lire en particulier |125, 126, 134 136|).

Pour toutes ces raisons, il est donc nécessaire et indispensable de travailler sur la matière brute : les codes source des virus. La certitude ne vient que par l'analyse du code. Là est la différence entre parler des virus et étudier les virus. Cette étude ne fera pas pour autant du lecteur un acteur malfaisant, bien au contraire. Chaque année, des milliers d'étudiants apprennent la chimie. Ils ne se mettent pas à fabriquer des armes chimiques à l'issue de leurs études. Doit-on proscrire l'enseignement de la chimie sous prétexte de risques relativement marginaux même s'ils sont particulièrement préoccupants ? Ce serait se priver de tout de ce que la chimie a apporté de bénéfique. Il en est de même pour la virologie informatique.

Une autre raison milite en faveur d'une présentation technique sur les virus. Les éditeurs d'antivirus, pour une grande partie d'entre eux, ont une part de responsabilité non négligeable dans la prolifération du risque viral. En effet, privilégiant la logique commerciale à travers un marketing souvent fallacieux, contestant aux autres le droit à la connaissance technique dans ce domaine, les utilisateurs en finissent par penser et par accepter qu'un antivirus est un outil de protection parfait, et que toute protection se réduit à disposer d'un tel produit. Il n'en est rien. Les utilisateurs doivent participer

activement à la lutte antivirale, à leur niveau. Cela n'est possible que s'ils disposent des connaissances adéquates.

Enfin, et cela milite en faveur de la nécessité d'une présentation technique de codes source viraux, il est nécessaire d'expliquer en prouvant, à l'aide de ce matériau brut, ce qui est possible ou non en matière de virus. Trop de décideurs fondent leur action ou leur prise de décision sur des concepts vagues et mal définis, relevant quelquefois du fantasme pur et simple. L'absence d'éléments techniques, pour séparer « le bon grain de l'ivraie », leur permet également de se conforter dans des certitudes lénifiantes mais dangereuses. Seule la confrontation avec la réalité effective d'un code source, élément de preuve irréfutable, permet d'envisager sainement les choses dans ce domaine.

Dans le présent ouvrages, les connaissances requises pour la bonne compréhension des notions exposées, ont été réduites au strict minimum. Une bonne connaissance des mathématiques de base, de la programmation ainsi que les rudiments de base concernant les systèmes d'exploitation Unix et Windows seront suffisants. L'optique de ce livre a été de privilégier ce que l'on pourrait appeler l'« algorithmique virale » et de montrer que les techniques virales peuvent être décrites simplement et indépendamment de tel ou tel langage ou de tel ou tel système d'exploitation (encore une fois cela replace la virologie informatique dans le contexte plus général de l'informatique et de la relation existant entre algorithmique et langages de programmation).

Dans ce but, l'usage du pseudo-code et du langage C a été systématiquement préféré quand cela était possible et pertinent. Ils sont généralement connus de la plupart des informaticiens. La présentation sera rendue plus aisée en considérant des exemples simples mais représentatifs, dans un langage accessible au plus grand nombre.

Certains lecteurs reprocheront, peut-être, de ne pas voir abordés en détails des pans entiers de la virologie informatique : les moteurs de mutation et le polymorphisme, les techniques avancées de furtivité ; et plus généralement de ne pas avoir consacré d'études aux virus et aux vers écrits en assembleur ou à l'aide de langages plus « exotiques » (mais importants ; citons Java, les langages de scripts type VBS ou Javascript, Perl, Postscript,...). Encore une fois, l'objet de cet ouvrage est une introduction didactique, pour le plus grand nombre, basée sur des exemples simples mais particulièrement représentatifs. Il est essentiel de comprendre l'algorithmique de base, commune à tous les virus et vers, avant de se polariser sur les spécificités de tel ou tel langage, de telle ou telle technique ou de tel ou tel système d'exploitation. Tous les aspects techniques évolués ou plus complexes de la virologie informatique, feront l'objet d'un second ouvrage faisant suite à celui-ci.

D'autres lecteurs pourront également reprocher à cet ouvrage de n'évoquer que très rapidement les techniques antivirales et de faire, en quelque sorte, la part belle aux seuls virus. En fait, cela n'est vrai qu'en apparence. La problématique de la sécurité (d'une manière très générale et non limitée au seul domaine informatique), est la situation nécessairement réactive à laquelle est condamné l'utilisateur. Les possibilités de détection d'une attaque, de protection et de prévention n'existent que par la connaissance que l'on a des actions offensives qui peuvent être menées. Dans le cas des virus, toute défense et toute lutte seront illusoires sans une connaissance claire et rigoureuse des mécanismes viraux.

Ce livre est articulé autour de trois parties, relativement indépendantes les unes des autres, que le lecteur pourra éventuellement consulter dans l'ordre qui lui plaira. Toutefois, la lecture préalable du chapitre 4 traitant de la taxonomie, des outils et des techniques de bases en virologie informatique est conseillée pour assimiler le vocabulaire de base et mieux comprendre le reste de l'ouvrage.

La première partie traite des aspects théoriques des virus. Les travaux de von Neuman sur les automates auto-replicatifs, de Kleene sur les fonctions récursives et de Turing sont présentés dans le chapitre 2. Ce sont les bases mathématiques indispensables pour la suite. Les formalisations de Fred Cohen et de Leonard Adleman sont ensuite exposées dans le chapitre 3. Elles sont fondamentales pour avoir une vision globale à la fois des virus et de la lutte antivirale. Sans cette hauteur, le lecteur passerait à côté de certains aspects et enjeux de la virologie informatique.

Le chapitre 4, ensuite, traite de la classification exhaustive des infections informatiques ainsi que des principales techniques et des outils. Il contient notamment les définitions essentielles qu'il convient de connaître pour la suite. Bien que sa lecture, préalable soit fortement conseillée, ce chapitre a été placé à cet endroit pour respecter la progression logique de l'ouvrage et l'historique du domaine. Le chapitre 5, enfin, clôt cette partie avec la présentation des techniques antivirales actuelles et du droit en matière de virologie informatique. Cette première partie pourra servir de support à une présentation théorique sur le sujet, de six heures environ. Que le lecteur non mathématicien se rassure. Chaque fois que cela a été possible, les concepts ont été simplifiés au maximum sans pour autant sacrifier la rigueur nécessaire.

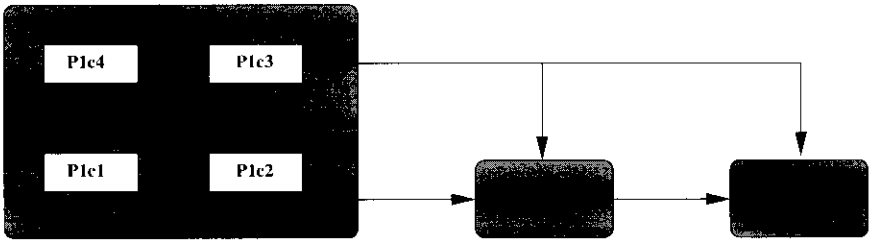
La seconde partie est nettement plus technique et consiste essentiellement en l'étude du code source de quelques virus représentatifs des principales familles. Là encore, le but a été de rendre ce livre accessible à des non-spécialistes et de limiter les prérequis nécessaires à une bonne connaissance

de la programmation. Seuls des virus assez simples mais très actuels et qui représentent encore une menace tout à fait réelle, sont considérés. Aussi pour cette première version, des techniques aussi passionnantes qu'ardues comme le polymorphisme ou la furtivité (et les techniques assimilées), ne seront pas traitées, autrement que d'un point du vue général. Ces techniques réclament des connaissances solides en assembleur. Le but principal de cet ouvrage est d'amener le lecteur à acquérir les connaissances qui lui permettront de poursuivre seul l'étude de la plupart des autres virus.

La troisième partie est peut être la plus importante des trois. Elle traite des applications des virus informatiques. Ces programmes particuliers sont potentiellement des outils puissants, aux nombreuses utilisations potentielles. Les rares ouvrages réellement techniques traitant des virus n'abordent pratiquement jamais cet aspect des virus. Considérant l'idée même virus " utiles " comme un début de remise en cause de leurs intérêts, les professionnels de la lutte antivirale l'ont frappé d'anathème. Il est autant absurde qu'illusoire et relève probablement d'une certaine forme d'obscurantisme, peut-être intéressé.

L'utilisation des virus à des fins applicatives n'est pas récente même si elle n'a pas été médiatisée. Maîtrisés comme il se doit (et les antivirus ont là un nouveau rôle, essentiel, à jouer, en les faisant évoluer de manière adéquate), les virus peuvent rendre de grands services. Cette partie tentera, à travers quelques exemples, de sensibiliser le lecteur.

Au final, l'articulation logique de cet ouvrage peut être résumée par le schéma suivant :



Ce livre reprend, en partie, le module de virologie informatique (entre 15 et 35 heures, travaux pratiques compris) dispensé à l'École Supérieure d'Électricité depuis 2002, à l'École Nationale Supérieure des Techniques Avancées depuis 2001, à l'École Spéciale Militaire de Saint-Cyr depuis 1999 et à l'université de Limoges depuis 2001. Il pourra aisément servir de support à tout enseignant désireux de monter un tel module. Chaque chapitre propose quelques exercices afin de permettre au lecteur désireux de poursuivre plus

avant la réflexion concernant les connaissances et techniques présentées. Dans certains cas, des ébauches de projet, d'une durée de deux à huit semaines, sont également proposées. Mon souhait est que ce livre fasse naître des initiatives pédagogiques permettant de faire découvrir les virus informatiques tout en les démystifiant.

Bien que ce livre représente, du moins je l'espère, un progrès appréciable dans la compréhension des virus informatiques, et devrait répondre à une demande qui ne fait que croître, je suis également conscient qu'il peut subsister, encore, quelques imperfections dans sa forme. Je prie les lecteurs de bien vouloir m'en excuser et je les invite d'ores et déjà à me faire part des éventuelles (mais inévitables, je le crains) coquilles et autres errata trouvées, afin d'améliorer progressivement cet ouvrage. Elles seront corrigées au fur et à mesure sur ma page web (`www-rocq.inria.fr/codes/Eric.Filiol/index.html`), page sur laquelle figureront également quelques corrigés d'exercices et autres informations utiles.

Ce livre est avant tout dédié à Fred Cohen. Sans lui, il est à craindre que la virologie informatique (les virus ET la lutte antivirale) serait encore une science balbutiante et immature. Son travail de formalisation et ses résultats n'ont malheureusement pas reçu l'attention qu'ils méritaient. Son apport est considérable et le but de ce livre est de contribuer, le mieux possible, à lui rendre hommage et à faire connaître une œuvre, à mon sens, majeure et incontournable.

Ce livre est également dédié à Mark Allen Ludwig, celui qui nous a ouvert la voie, à tous, en publiant quelques livres techniques sur les virus, avec de nombreux codes source détaillés. Son approche didactique, intelligente, éclairée (le mot n'est pas trop fort) autant que constructive et objective est remarquable. La rigueur scientifique avec laquelle il s'est attaché à décrire des techniques avant tout passionnantes et stimulantes pour l'esprit, son œuvre dans de domaine, tient à ce jour en quatre livres dont la lecture reste incontournable en a fait un pionnier. Mark Ludwig ne renierait pas lui-même ce terme qui lui est cher. Il est devenu en quelque sorte un guide pour tous les passionnés de virus et d'intelligence artificielle. Beaucoup de personnes lui doivent cette passion quasi-naturaliste pour les virus informatiques. Je ne saurais non plus oublier Pascal Lointier, président du CLUSIF, qui en assurant la traduction d'un des livres majeurs de Mark Ludwig, a permis aux lecteurs francophones d'accéder à un livre passionnant. Il a lui-même grandement contribué en France à donner une vision saine et pédagogique de la virologie informatique. Nombreux sont ceux en France, qui lui doivent beaucoup.

En troisième lieu, je souhaiterais dédier ce livre à certains programmeurs de virus, le plus souvent anonymes (sauf pour les initiés) mais assurément géniaux, animés par le sens du défi technique et non par une quelconque envie de nuire. Le code de certains de leurs virus m'a émerveillé, a alimenté cette passion pour les virus et au-delà pour le génie humain qui ne se satisfait d'aucune impossibilité technique. Leur apport est considérable, plus que l'on ne veut bien le dire en général. Ils m'ont, en particulier, convaincu que dans le domaine de la virologie informatique (mais cela est vrai dans tous les domaines du savoir), la principale qualité est l'humilité. Il ne faut pas se complaire du peu que l'on a pu apprendre mais toujours regarder la masse impressionnante et insolente de ce que l'on ignore. Trop de gens se prétendent experts mais ignorent que les techniques évoluent.

Enfin, je tiens à remercier, pour des raisons diverses, les personnes qui ont rendu ce livre possible : avant tout Madame Huilleret et Monsieur Puech des éditions Springer, qui ont immédiatement été séduits par ce projet, les lieutenants Azatassou, De Gouvion Saint-Cyr, Hélo, Plan, Smitsomboon, Tanakwang, les lieutenants de vaisseau Ratier et Turcat, qui ont participé aux développement de certaines versions des virus, lors de leur stage d'ingénieur, au sein du laboratoire de virologie et de cryptologie de l'École Supérieure et d'Application des Transmissions, mais également le général Desvignes et les lieutenant-colonels Gardin et Rossa qui, à leur façon, m'ont soutenu dans cette entreprise et ont compris la nécessité de développer, chez les futurs professionnels de la Défense, une véritable culture technique en matière de virologie informatique ; Christophe Bidan, Nicolas Brulez, Jean-Luc Casey, Thiébaut Devergranne, le chef d'escadron Alain Foucal, Brigitte Jülg, Pierre Loidreau, Marc Maiffret, Thierry Martineau, Arnaud Metzler, Bruno Petazzoni, Frédéric Raynal, Eugène H. Spafford, Denis Tatania, Alain Valet, pour m'avoir permis de faire partager à d'autres cette passion, tous mes élèves et étudiants sans lequels les cursus créés n'auraient pas vu le jour. Enfin, ma femme Laurence qui a réalisé le CDROM fourni avec cet ouvrage et m'a soutenu avec patience dans cette entreprise.

Maintenant entrons dans le monde fantastique des virus informatiques et de l'algorithmique virale.

Guer, août 2003,

Éric Filiol
Eric.Filiol@inria.fr

Table des matières

Conclusion