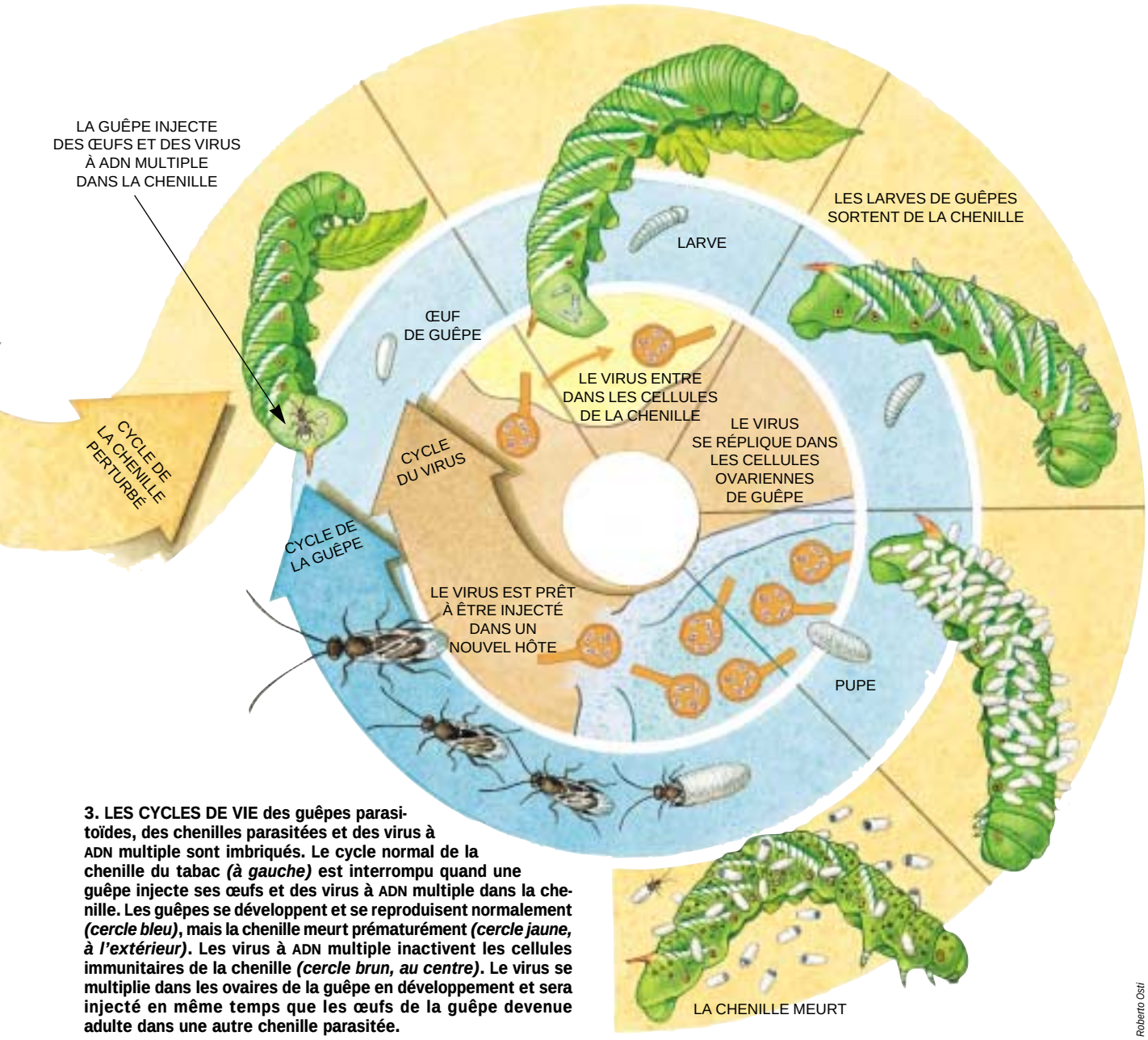


3. LES CYCLES DE VIE des guêpes parasitoïdes, des chenilles parasitées et des virus à ADN multiple sont imbriqués. Le cycle normal de la chenille du tabac *(à gauche)* est interrompu quand une guêpe injecte ses œufs et des virus à ADN multiple dans la chenille. Les guêpes se développent et se reproduisent normalement *(cercle bleu)*, mais la chenille meurt prématurément *(cercle jaune, à l'extérieur)*. Les virus à ADN multiple inactivent les cellules immunitaires de la chenille *(cercle brun, au centre)*. Le virus se multiplie dans les ovaires de la guêpe en développement et sera injecté en même temps que les œufs de la guêpe devenue adulte dans une autre chenille parasitée.

Roberto Osti

gènes de venin qui leur auraient été utiles. Selon une autre hypothèse, les guêpes auraient trouvé un moyen efficace d'utiliser leurs propres protéines de venin en recopiant leurs gènes, en les regroupant et en les transférant dans les cellules de chenille où elles continuent à agir. Dans les deux cas, la synergie entre les guêpes et les virus augmente, et l'origine de la caractéristique est effacée.

Quelle que soit l'origine des virus à ADN multiple, leur association avec les guêpes et les chenilles permet aux biologistes de préciser les mécanismes de l'évolution. La redoutable tactique des guêpes parasitoïdes et de leurs complices viraux contredit l'hypothèse selon laquelle les parasites les plus évolués sont peu virulents vis-à-vis de leurs hôtes : les guêpes parasitoïdes tuent tous leurs hôtes sans exception, en suivant un scénario bien réglé où elles gagnent toujours. À l'inverse, les guêpes et les virus à ADN multiple entretiennent une association mutuellement bénéfique si étroite que les frontières génétiques entre ces deux espèces s'estompent. Pourquoi la chenille ne devient-elle pas un papillon? La question reste ouverte et continue d'intriguer les biologistes de l'évolution, les entomologistes et les endocrinologues.

Nancy BECKAGE est professeur d'entomologie à l'Université de Riverside.

Jo-Ann G.W. FLEMING, *Polydnaviruses : Mutualists and Pathogens*, in *Annual Review of Entomology*, vol. 37, pp. 401-426, 1992.

James H. TUMLINSON, W. Joe LEWIS et Louise E.M. VET, *Comment certaines guêpes détectent leurs hôtes*, in *Pour la Science*, mai 1993.

M.D. LAVINE et N.E. BECKAGE, *Polydnaviruses : Potent Mediators of Host Insect Immune Dysfonction*, in *Parasitology Today*, vol. 11, n° 10, pp. 368-378, 1995.

Donald L.J. QUICKE, *Parasitic Wasps*, Chapman & Hall, 1997.

J. van BAAREN, G. BOIVIN et J.-P. NÉNON, *Guêpes et apprentissage*, in *Pour la Science*, août 1997.

Comment combattre les virus informatiques

JEFFREY KEPHART • GREGORY SORKIN • DAVID CHESS • STEVE WHITE

La biologie éclaire le comportement des virus informatiques et aide à concevoir des parades.

Naguère, les épidémies redoutées étaient la peste, le choléra ou la grippe espagnole. À l'ère des autoroutes de l'information, les œuvres de fiction grouillent de virus informatiques qui se dupliquent et détruisent les ordinateurs à travers les réseaux. Si les virus infectent effectivement un nombre croissant d'ordinateurs, ils ne sont pas la menace que dépeignent les auteurs de science-fiction. Est-ce pour cette raison qu'ils n'ont pas encore été beaucoup étudiés par les informaticiens?

Leur notoriété est due à une particularité qu'ils partagent avec les véritables virus : ils se lient à un système dont ils utilisent les ressources pour se répliquer (les virus biologiques se lient à une cellule, les virus informatiques à un programme ou à un ordinateur). Leurs effets sont variables : certains virus informatiques sont inactifs, d'autres gênants, et d'autres encore désastreux. À la manière des virus biologiques qui infectent les individus d'une communauté, les virus informatiques se déplacent de programme en programme, d'ordinateur en ordinateur

Avec mes collègues du Centre de recherche *IBM* de Yorktown Heights nous avons trouvé qu'une analogie avec la biologie aide à comprendre les épidémies de ces virus et à élaborer des défenses. En utilisant les outils mathématiques mis au point ces dernières décennies par les épidémiologistes, nous avons compris certains facteurs qui gouvernent leur vitesse de propagation. Nos découvertes de méthodes de détection efficaces de ces virus et des relations entre virus doivent beaucoup aux techniques de recherche d'homologie mises au point en biologie moléculaire. De surcroît, nous avons trouvé des parades contre ces logiciels dangereux en nous inspirant du système immunitaire des vertébrés et de sa remarquable capacité à repousser et à détruire des agents pathogènes.

Le plus ancien ancêtre des virus informatiques n'a existé que théoriquement : en 1940, John von Neumann étudie l'autoréplication des automates cellulaire. Si l'idée de programmes capables d'infecter des ordinateurs date de 1970, le premier cas d'infection répertorié date du mois d'octobre 1987, lorsqu'un petit programme nommé *Brain* («cerveau») apparut sur plusieurs dizaines de disquettes à l'Université du Delaware. Aujourd'hui, les virus infectent au moins un million d'ordinateurs par an, et les utilisateurs dépensent chaque année plusieurs centaines de millions de francs en réparations et en logiciels antiviraux.

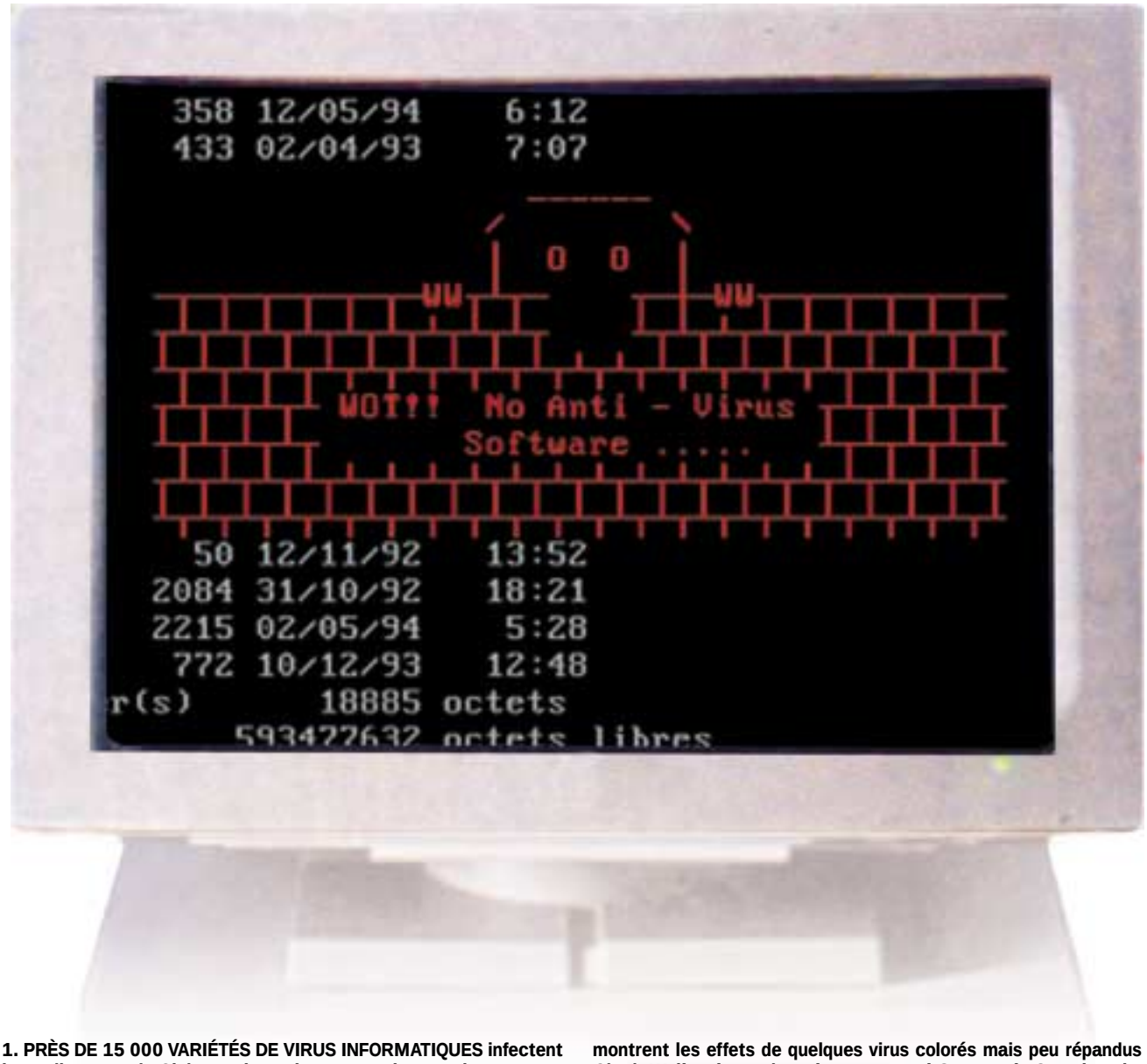
La plupart des virus s'attaquent à des micro-ordinateurs de type PC, c'est-à-dire des compatibles avec les ordinateurs *IBM*. Près de 15 000 virus ont déjà vu le jour, et six nouveaux virus apparaissent quotidiennement. Heureusement, seuls quelques-uns d'entre eux se sont disséminés.

Il existe trois types principaux de virus : les virus qui contaminent les programmes exécutables, les virus de secteur de démarrage et les virus de macros. Environ 85 pour cent des souches virales contaminent des programmes, comme des tableurs ou des jeux, mais cela ne représente que 20 pour cent des infections. Lorsque l'utilisateur exécute un programme infecté, les lignes du programme viral s'exécutent en premier et se copient dans la mémoire vive de l'ordinateur, de façon à se recopier toutes seules, dans un autre programme qui sera lancé plus tard. Une fois installé dans la mémoire, le virus redonne le contrôle au programme infecté, si bien que l'utilisateur ne peut déceler sa présence. Lorsque le programme infecté atteint un autre ordinateur par l'intermédiaire d'une disquette ou d'un réseau, le cycle d'infection recommence.

Les virus de secteur de démarrage représentent cinq pour cent de l'ensemble des virus connus. Ils résident dans une zone des disquettes ou des disques durs qui est lue quand l'ordi-

nateur est allumé et dont les programmes sont exécutés automatiquement. Habituellement, le secteur de démarrage contient les instructions de chargement du système d'exploitation, le système qui gère le fonctionnement de l'ordinateur et qui permet notamment l'affichage des données ou la lecture d'ordres tapés par l'utilisateur. Une fois installé, un tel virus peut contaminer toute disquette non protégée en écriture introduite dans le lecteur. Il infecte aussi le disque dur, de sorte qu'il se trouve chargé dans la mémoire vive de l'ordinateur à chaque redémarrage. Les virus de démarrage sont très efficaces et, bien que moins de 1 000 souches différentes existent, ils sont pour le moment plus répandus que les virus qui infectent les programmes exécutables (30 pour cent des infections).

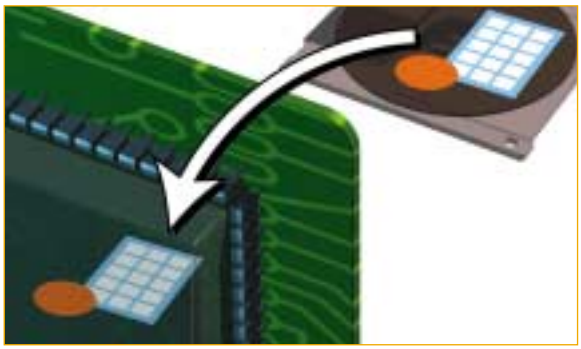
La troisième catégorie, les virus de macros, sont indépendants des systèmes d'exploitation, de sorte qu'ils infectent aussi bien des ordinateurs compatibles PC que des *Macintosh* ; ils infectent des fichiers de données plutôt que des programmes. Beaucoup de tableurs, de programmes de bases de données et de traitement de texte peuvent exécuter des séquences de commandes nommées «macros», qui automatisent des tâches fastidieuses : certaines macros de traitement de texte placent automatiquement un mot long ou une expression toute entière quand l'utilisateur tape seulement quelques lettres ; des macros de tableurs effectuent de longs calculs. Des concepteurs de virus ont programmé des macros qui se recopient dans d'autres documents. De tels virus se propagent beaucoup plus vite que les autres formes de virus, parce que de nombreuses personnes s'échangent des fichiers de données, lorsque, par exemple, ils rédigent un document en commun. Le premier virus de macros détecté, un virus nommé *Concept*, fut trouvé fin 1995 dans un document créé par le programme de traitement de texte *Word*, et c'est aujourd'hui l'un des virus les plus répandus (dix pour cent des infections), avec un nouvel arrivant, *CAP.A*, qui représente 20 pour cent des infections. En tout, on a recensé plus de 1 500



Jana Brenning

1. PRÈS DE 15 000 VARIÉTÉS DE VIRUS INFORMATIQUES infectent les ordinateurs du Globe, mais quelques centaines seulement sont répandues. Les écrans reproduits ci-dessus et sur la page ci-contre montrent les effets de quelques virus colorés mais peu répandus : *Chad*, *Walker* («marcheur») et *Rescue* («Secours»). Certains virus sont inoffensifs, mais d'autres détruisent irrémédiablement les don-

Le cycle d'un virus qui infecte les programmes



Lorsqu'un utilisateur lance un programme infecté, l'ordinateur copie le programme du disque, où il est stocké et inactif, dans la mémoire vive, où il peut être exécuté.

Le programme viral s'exécute en premier, tandis que le programme infecté reste au repos.

Le virus se recopie dans une zone de la mémoire vive séparée du programme, de sorte qu'il peut rester actif même après que l'utilisateur a lancé d'autres logiciels.

virus de macros, responsables de la moitié de toutes les attaques virales.

En plus des séquences qui assurent la recopie, les virus contiennent des séquences variées : certains virus impriment un message ou affichent une image ; d'autres effacent des programmes ou des données. Même ceux qui ne sont potentiellement pas nuisibles peuvent créer des dégâts quand ils se placent dans des configurations différentes de celle où ils avaient été créés. Par exemple, le virus *Form*, qui ne produit habituellement qu'un léger cliquetis chaque mois, écrit sur une partie du disque qui est inoccupée dans les vieux ordinateurs, mais qui contient des données cruciales dans les nouvelles machines, où les répertoires sont agencés différemment.

Techniques immunitaires

Les logiciels contre les virus sont apparus peu après les virus. Des programmes généraux de détection de virus surveillent si le comportement d'un ordinateur est modifié par la présence d'un virus (par exemple, si le virus modifie certains fichiers importants ou des parties de la mémoire) et cherchent périodiquement des modifications suspectes des programmes. Ils détectent les virus sans les connaître, mais ils peuvent aussi déclencher de fausses alarmes, car des activités normales de l'ordinateur ressemblent parfois à celles de virus en action.

En revanche, des programmes d'analyse recherchent dans les fichiers, dans les secteurs de démarrage ou dans la mémoire de l'ordinateur, des morceaux de programmes (nommés aussi signatures virales) caractéristiques de virus déjà connus. Pour rester efficaces, ils doivent être mis à jour régulièrement, mais ils ne déclenchent que rarement de fausses alarmes. Ces programmes ont la même efficacité que les récepteurs biologiques du système immunitaire : ces derniers reconnaissent les protéines virales en se liant à de courtes séquences (8 à 15 acides aminés seulement, contre des milliers pour la protéine virale complète) ; de même, les signatures virales reconnues par les programmes d'analyse sont très courtes, de 16 à 30 octets (un octet représente huit bits), parmi les quelques milliers qui constituent un virus informatique complet. La recherche de petits segments, au lieu du virus complet, accélère l'analyse, d'autant que des virus différents partagent parfois la même signature. La plupart des programmes d'analyse utilisent des algorithmes issus des techniques d'appariement en biologie, qui cherchent plusieurs signatures en même temps. Le plus rapide de ces programmes vérifie la présence de 10 000 signatures dans 10 000 programmes en moins de dix minutes.

Comment éliminer un virus qui a été détecté? Une technique brutale, mais efficace, consiste à effacer le programme contaminé. La technique est identique à celle qui est utilisée par certains types de cellules immunitaires, qui détruisent les cellules infectées ; mais si les cellules du corps sont aisément remplacées, il n'en est pas de même des programmes et des documents informatiques. Les programmes antiviraux perfectionnés essaient plutôt de réparer les fichiers infectés. Cette tâche est facilitée par le fait que les virus informatiques ne se transmettent que s'ils n'éveillent pas l'attention de l'utilisateur, de sorte qu'ils laissent leur programme hôte intact.

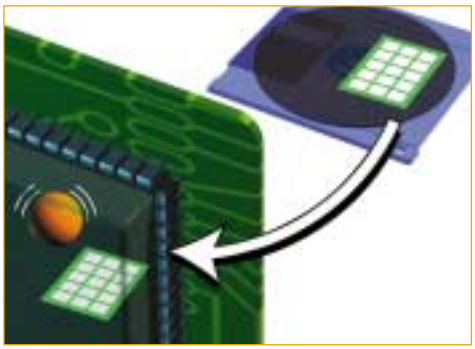
Lorsqu'un programme d'analyse détecte un fichier infecté, il exécute habituellement une ordonnance détaillée, fournie par ses concepteurs, afin d'ôter le code viral et de reconstruire une copie opérationnelle du programme original. D'autres techniques de désinfection générale fonctionnent aussi bien sur des virus connus que sur des virus inconnus. Nous avons mis au point un programme qui élabore un résumé mathématique (nommé somme de contrôle) de chacun des programmes du système, de sorte que l'on peut reconstituer une copie de l'original lorsqu'un programme est infecté.

Les techniques de détection et d'éli-

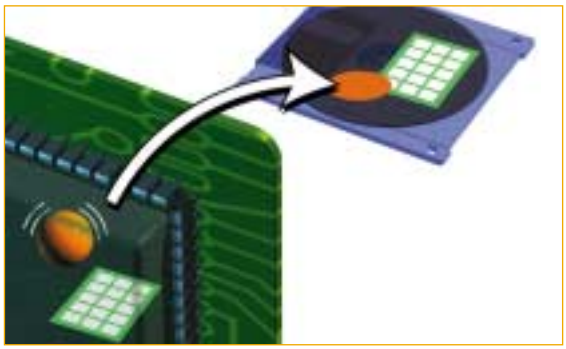
James Gary

2. SUR LE CHAMP DE BATAILLE NUMÉRIQUE, des programmes antiviraux recherchent les programmes malfaisants (virus, vers ou chevaux de Troie) et réparent les dégâts qu'ils causent. Les nombreux types de virus informatiques qui existent peuvent être reconnus à leur action et à leur structure caractéristique. Les programmes antiviraux protègent aussi les programmes ordinaires en les analysant et en prévenant les utilisateurs de tout changement.

Son travail initial effectué, le virus redonne le contrôle au programme infecté.



Lorsque l'utilisateur lance un autre programme, le virus dormant se réveille.



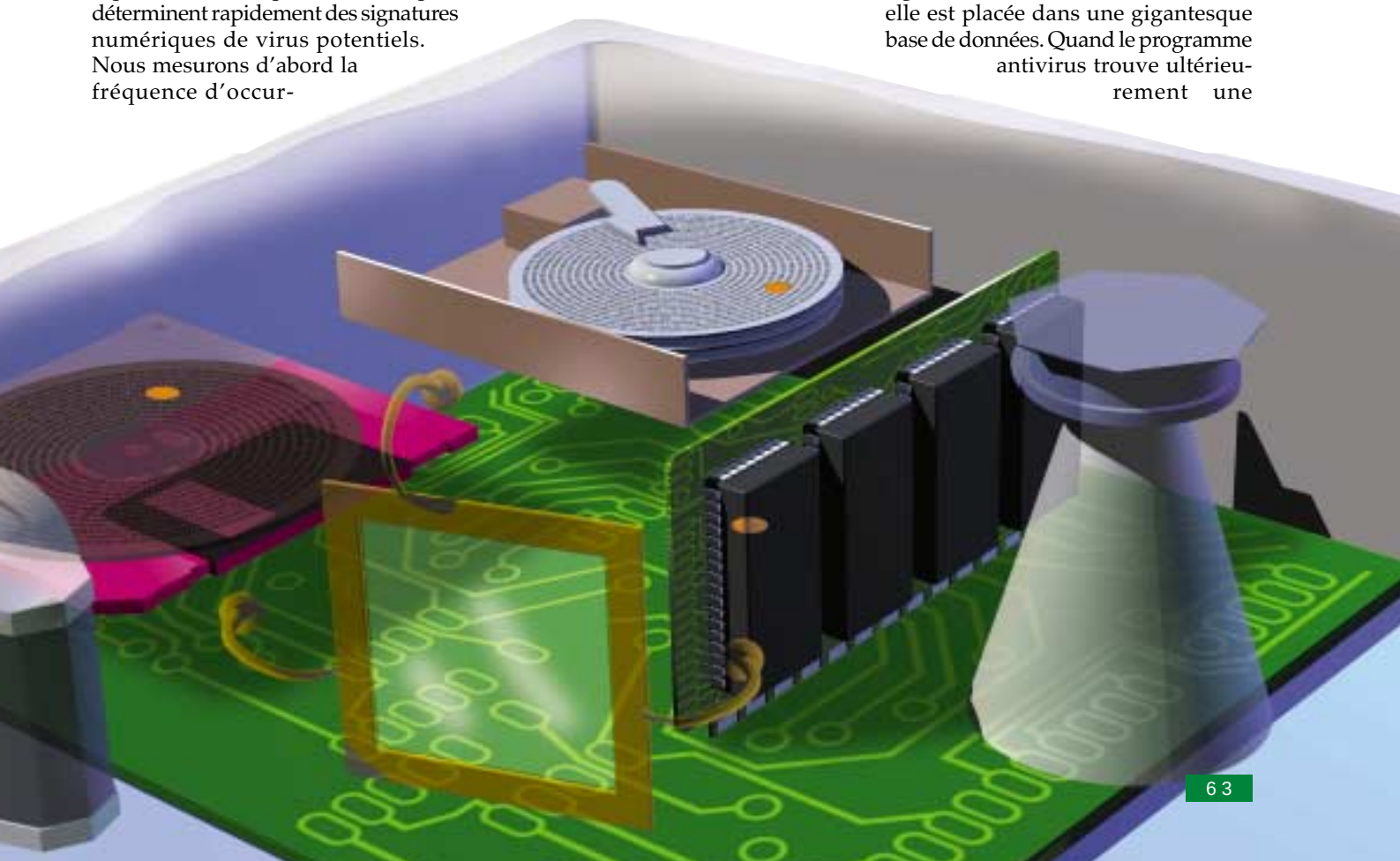
Il se copie dans le logiciel sain et peut aller infecter un autre ordinateur ou une disquette.

James Gary

mination de virus spécifiques nécessitent une analyse détaillée de chaque nouveau virus. Des spécialistes doivent identifier des séquences inhabituelles d'instructions qui apparaissent dans le programme viral, mais pas dans les programmes normaux. Ce processus est fondé sur une bonne connaissance des virus, mais aussi sur l'intuition. Ces spécialistes élaborent également les prescriptions pour enlever le virus de n'importe quel ordinateur hôte. Comme six nouveaux virus sont créés chaque jour, nous avons mis au point des outils et des procédures automatiques, afin d'aider les spécialistes humains, voire de les remplacer.

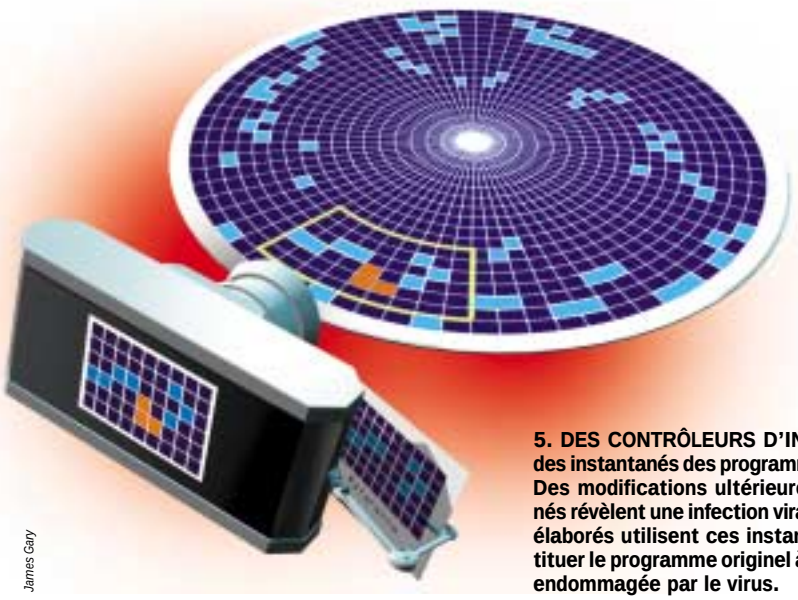
Fondées sur des méthodes statistiques, ces techniques automatiques déterminent rapidement des signatures numériques de virus potentiels. Nous mesurons d'abord la fréquence d'occurrence de certaines courtes séquences d'octets dans un grand nombre de programmes normaux. Quand on nous envoie un nouveau virus, notre logiciel le détecte en trouvant la séquence qu'il est statistiquement le moins probable de rencontrer dans les programmes légitimes. Cette méthode est beaucoup plus rapide qu'une analyse manuelle, et les signatures identifiées conduisent à moins de fausses alarmes que les signatures choisies par des experts humains. Cette méthode heuristique, initialement mise au point par Fridrik Skulason et Franz Veldman, ressemble à un mécanisme qui avait semblé être celui du système immunitaire : on croyait naguère que les anticorps se modelaient d'après les envahisseurs ; de même, nos signatures sont réalisées spécifiquement pour chacun des nouveaux virus rencontrés.

Avec ses collègues, Stephanie Forrest a mis au point une autre méthode, copiée sur la «sélection clonale», qui semble être le mécanisme du système immunitaire : le corps produit d'emblée une grande variété de cellules immunitaires, mais seules prolifèrent les cellules qui reconnaissent l'agent pathogène. De même, dans le programme fondé sur cette méthode, des signatures sont engendrées au hasard, et chaque signature est ensuite comparée à tous les programmes de l'ordinateur non infecté ; quand une signature ne ressemble à rien de connu, elle est placée dans une gigantesque base de données. Quand le programme antivirus trouve ultérieurement une



3. UN PROGRAMME ANTIVIRAL GÉNÉRAL surveille les comportements qui peuvent être dus à un virus en action, telle la modification de certaines zones de la mémoire vive ou de fichiers particuliers du disque dur. En bloquant ces comportements illicites, il n'élimine pas le virus, mais peut l'empêcher d'infecter d'autres programmes ou de perturber le fonctionnement de l'ordinateur.

4. UN ANALYSEUR DE SIGNATURES recherche sur les disques des fragments de programme qui apparaissent dans les virus connus.



5. DES CONTRÔLEURS D'INTÉGRITÉ prennent des instantanés des programmes et des données. Des modifications ultérieures de ces instantanés révèlent une infection virale. Des algorithmes élaborés utilisent ces instantanés pour reconstituer le programme originel à partir de la version endommagée par le virus.

James Gary

telle signature dans un programme, il sait que le programme a été modifié, bien qu'une analyse ultérieure soit requise pour déterminer si un virus en est la cause.

La biologie est également à la base d'une autre technique antivirale : on recherche des parentés entre les nouveaux et les anciens virus. En effet, les programmeurs de virus utilisent souvent des parties de virus existants pour en créer de nouveaux. Ces parties, ou «gènes» viraux, permettent de retracer l'évolution historique des virus informatiques, comme le font les biologistes lorsqu'ils établissent les relations entre espèces vivantes. En étudiant de nombreux virus, nous construisons automatiquement un ensemble de signatures valables pour tous les membres de la famille virale, incluant des variantes précédemment inconnues. Cette technique réduit les problèmes de stockage des signatures : une seule signature de 20 octets peut reconnaître des dizaines de virus différents.

Nous avons également mis au point des techniques pour reconnaître des virus en analysant la présence de très courts segments de programme, chacun de trois à cinq octets de longueur. Ces segments minuscules représentent des instructions dont l'exécution est spécifique d'une infection virale. Bien que des logiciels classiques puissent en contenir certains, leur répétition témoigne d'une infection virale. Les logiciels antiviraux en détectent la présence et, ce qui est plus important, reconnaissent de nombreux virus sans les avoir rencontrés auparavant.

La propagation des virus en milieu naturel

Depuis 1990, nous avons étudié les virus sur une population de plusieurs centaines de milliers d'ordinateurs qui nous avaient été confiés par des clients. Nous enregistrons le lieu et la date de chaque incident, le nombre d'ordinateurs ou de disquettes infectés et l'identité des virus. Ces statistiques nous ont permis de comprendre la propagation des infections virales dans le milieu naturel. Seulement cinq pour cent des virus connus ont été observés dans notre échantillon, et certains n'ont été vus qu'une seule fois. Les dix virus les plus fréquents déclenchent deux tiers des incidents, et leur prévalence est semblable : au cours du temps, le nombre de ces virus à succès augmente

linéairement jusqu'à atteindre un plateau ; ensuite, ils continuent d'apparaître à un rythme constant dans les ordinateurs, et quelquefois ils disparaissent complètement.

Afin de comprendre ces caractéristiques, nous avons utilisé des modèles mathématiques élaborés par les épidémiologistes. Les modèles les plus simples prédisent l'évolution d'une infection à partir d'un petit nombre de paramètres : le «taux de naissance», c'est-à-dire la vitesse à laquelle les individus malades infectent les autres, et le «taux de mortalité», la vitesse à laquelle un individu malade meurt ou est guéri. Lorsque le quotient de ces deux quantités est inférieur à un seuil, l'infection disparaît rapidement. Plus ce quotient est élevé et plus, en l'absence de mesures prophylactiques ou de réactions immunitaires, le nombre de personnes infectées augmente rapidement.

Toutefois, cette théorie simple ne rend pas bien compte de l'évolution des virus informatiques, car elle stipule que, sauf quand le quotient des naissances et des morts est proche de la valeur critique, les virus disparaissent ou se propagent à toute la population. Or, de nombreux virus infectent une très faible proportion de la population et restent à ce niveau. Pourquoi cette évolution? Parce que la théorie suppose qu'un individu a la même chance de contaminer n'importe quel autre individu de la population à risque. En informatique, ce n'est pas le cas, et l'on doit compliquer la théorie pour décrire la structure en groupes fermés de la population qui utilise les ordinateurs : en moyenne, chaque personne ne partage des logiciels et des données qu'avec un petit nombre de personnes, et tous les échanges prennent place à l'intérieur de groupes. Si Philippe échange des logiciels avec Pascale et Pascale avec Françoise, alors, il est raisonnable de penser que Philippe et Françoise peuvent échanger des virus.

Des simulations informatiques ont montré que la proximité des contacts ralentit la croissance des épidémies conformément à nos observations. Des échanges peu fréquents réduisent la probabilité d'une épidémie et abaissent le niveau du plateau, mais n'expliquent pas complètement les données.

L'évolution en action

De même que les disettes, les campagnes de santé publique ou les migrations modifient le cours des épidémies biologiques, des changements de l'environnement informatique sont responsables de plusieurs époques distinctes dans l'histoire de la virologie informatique. Jusqu'en 1992, les virus infectant les programmes ou les secteurs de démarrage se trouvaient

Les virus de messagerie électronique

La croissance rapide du réseau Internet accroît proportionnellement les menaces liées au contenu des messageries électroniques, et plus particulièrement, aux virus informatiques. Un message électronique qui transite sur le réseau peut être séparé en trois parties : l'en-tête, le corps du message et les fichiers attachés. La principale cible des virus informatiques est le fichier attaché. Ces fichiers sont soit des programmes exécutables, soit des documents. Aujourd'hui, on se prémunit contre ce type d'infection informatique par deux approches principales : la protection répartie et la protection centralisée.

Dans la protection répartie, chaque poste susceptible de recevoir un message est protégé. La plupart des programmes antiviraux du marché disposent de modules de surveillance. Ces derniers détectent la présence du virus lorsque les éléments attachés sont exécutés ou ouverts. Malheureusement, cette méthode n'empêche pas la propagation du virus sur l'ensemble des postes des destinataires de messages : si, sur certains postes infectés, le virus n'est pas éradiqué, l'épidémie peut s'étendre. De surcroît, la détection n'intervient qu'à l'ouverture du fichier. Dès lors, un fichier attaché infecté pourra être transmis d'un utilisateur à l'autre sans qu'aucune alerte soit donnée.

Certains programmes antiviraux possèdent des modules spécifiques pour la messagerie électronique. Ces programmes décomposent le message et l'analysent sans ouvrir ni exécuter le fichier attaché. Si cette détection précoce évite l'épidémie, elle dépend du type de messagerie électronique et n'empêche pas la propagation d'un virus infecté à de nombreux utilisateurs. Ainsi, une note de service envoyée à l'ensemble du personnel d'une société entraînera de nombreuses alertes.

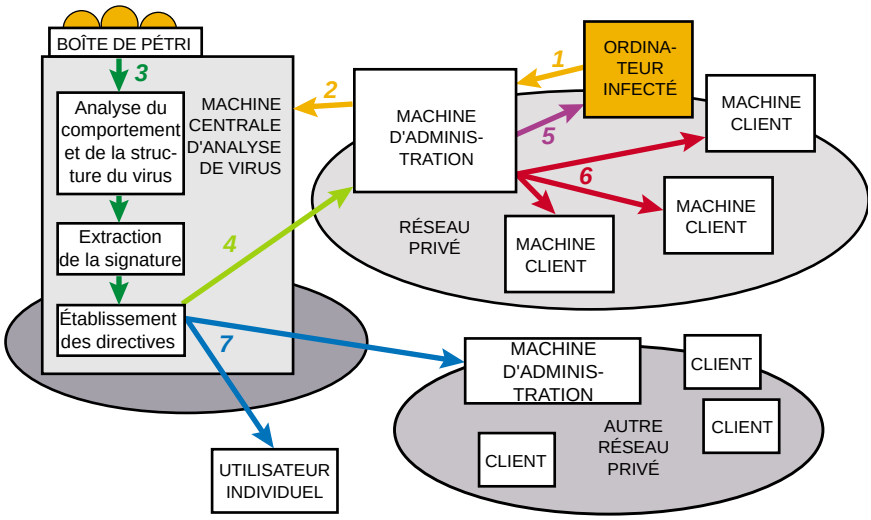
L'autre approche, nommée protection centralisée, consiste à utiliser une machine de contrôle. Tous les messages qui transitent passent par cette machine. À leur arrivée, les messages sont décomposés et décompressés (les fichiers attachés sont souvent compressés). Sur la machine centrale, des logiciels antivirus éradiquent alors les éventuels virus avant de faire suivre le message. Cette solution a l'avantage d'être préventive et centralisée : la note de service infectée envoyée à l'ensemble du personnel sera bloquée avant d'atteindre ses destinataires. Malheureusement, ces solutions ne fonctionnent qu'avec certains types de messageries.

La dernière approche qui s'apparente également à de la protection centralisée, n'est pas préventive : elle consiste à analyser tous les messages présents sur une machine centrale. Cette machine centrale, ou serveur de messagerie, vérifie l'innocuité des messages avant qu'ils ne parviennent à l'utilisateur. Cette solution est encore peu répandue, mais elle présente l'intérêt d'être économique en terme de ressource mémoire et de système.

Quelle que soit la solution mise en œuvre, l'utilisation de l'Internet et des messageries électroniques doit se faire avec une extrême prudence. Sans protection adéquate, quelques règles élémentaires de conduite peuvent éviter le pire : ne pas ouvrir un message en provenance d'un expéditeur inconnu, ne pas exécuter de programmes ni ouvrir de document attaché sans en connaître l'origine et le but, sauvegarder sur disquette les zones et les fichiers sensibles.

Sans ces précautions, le *surf and buy* («se promener sur l'Internet et acheter») peut rapidement se transformer par un *surf and die* («se promener» sur l'Internet et mourir) pour les disques durs et les informations des internautes non avertis.

Stéphane BOUCHÉ, Société *OMNISET*



6. DANS LES RÉSEAUX, des systèmes immunitaires informatiques empêcheront la prolifération des virus. Un ordinateur infecté envoie une copie du programme suspect à une machine d'administration *(1)* qui, à son tour, envoie ce programme, sous forme codée, à une machine centrale d'analyse de virus *(2)*. Cette machine cultive le virus dans une «boîte de Pétri» informatique, où il analyse son fonctionnement et sa structure *(3)*. Les recommandations qui en résultent sont renvoyées vers la machine d'administration *(4)*, qui les renvoie tout d'abord vers l'ordinateur infecté *(5)*, puis vers les autres machines du réseau local *(6)*. Les autres abonnés du monde entier reçoivent périodiquement des remises à jour d'antiviraux qui les protègent des nouveaux virus *(7)*.

en quantités égales, et leur nombre croissait légèrement. Puis, les virus de programmes sont devenus moins fréquents, tandis que le nombre de virus de secteur de démarrage continuait à augmenter. Entre fin 1992 et fin 1995, ces derniers régnèrent sans partage. Pourquoi les virus de programmes se sont-ils éteints?

Nous supposons qu'ils ont été bloqués par l'avènement du système d'exploitation *Windows 3.1*, une amélioration du système MS-DOS utilisé sur les ordinateurs compatibles PC à partir de 1992 : en présence de certains virus qui infectent les programmes, le système *Windows 3.1* se bloque, de sorte que les utilisateurs se sont débarrassés des virus (parfois en vidant le contenu du disque dur et en installant à nouveau tous les logiciels), même quand ils ignoraient que leurs problèmes provenaient d'un virus. En revanche, les virus de démarrage coexistent pacifiquement avec *Windows 3.1*, ils ne gênent pas leur hôte avant que l'infection ne se déchaîne.

L'utilisation générale de *Windows 95*, un autre système d'exploitation, a provoqué le brusque déclin des virus de démarrage. Ce système d'exploitation prévient l'utilisateur de la plupart des changements survenant dans la zone de démarrage, si bien que les virus ne peuvent s'y développer. Lors de notre recensement, nous avons observé quelques virus de démarrage spécifiques à *Windows 95* et à d'autres systèmes d'exploitation à 32 bits, mais ils ont peu de chances de se multiplier.

Aujourd'hui, l'ère des virus de macros a commencé. Ces virus, attachés aux fichiers de données, sont échangés plus fréquemment que les virus attachés aux programmes : ils se transmettent plus rapidement que les autres virus. À l'heure des réseaux, des programmes de courrier électronique et de transfert de fichiers élaborés permettent de partager des programmes et des documents plus rapidement et plus facilement que naguère, ce qui amplifie le problème.

Les virus de macros sont aussi les premiers qui exploitent la tendance croissante à l'interopérabilité des ordinateurs : un virus qui infecte des fichiers DOS ne met pas en danger un *Macintosh*, par exemple, parce que les instructions de copie des systèmes DOS diffèrent de celles des *Macintosh* ; en revanche, un virus de macros peut infecter n'importe quel ordinateur du moment qu'il possède un programme vulnérable. Le fait que le traitement de texte *Word* fonctionne sur une grande variété d'ordinateurs permet à *Concept*, ainsi qu'à d'autres virus de macros, de franchir les frontières entre systèmes.

Aujourd'hui, les virus se déplacent d'un ordinateur à un autre quand les utilisateurs s'échangent volontairement et manuellement des disquettes ou d'autres supports numériques ; on parvient à enrayer les épidémies, car même les virus les plus efficaces mettent des mois, voire des années, pour se propager. Dans les réseaux, les virus se propagent beaucoup plus vite. En 1988 déjà, Robert Tappan Morris avait lancé le «vers de l'Internet», un programme qui exploitait les défauts de sécurité du système d'exploitation UNIX et qui a envahi des centaines d'ordinateurs sur toute la Terre en moins d'une journée.

Un système immunitaire numérique

Les nouvelles techniques de chargement automatique des programmes et des données aggravent le problème. Les programmes de courrier électronique permettent l'envoi de textes ou de feuilles de calcul à des messages : lorsque l'on ouvre le document attaché, le programme adapté s'ouvre, et tous les virus de macros contenus dans le document sont exécutés. Plus grave encore, des «agents intelligents», c'est-à-dire des programmes qui parcourent le réseau à la place des utilisateurs, enverront périodiquement des courriers électroniques et ouvriront automatiquement les documents attachés ; les virus se propageront à des vitesses bien supérieures à celles d'aujourd'hui.

Ces changements de l'écosystème numérique imposent la recherche rapide d'une parade automatique aux problèmes posés par les virus. Les Sociétés *IBM*, *Symantec* et *McAfee* mettent au point des techniques d'analyse et d'éradication automatiques des nouveaux virus. Chez *IBM*, nous créons une sorte de système immunitaire pour les réseaux. De même que le système immunitaire des vertébrés crée des cellules qui reconnaissent et combattent de nouveaux agents pathogènes après quelques jours d'exposition, un système immunitaire informatique peut donner, en quelques minutes, des directives pour reconnaître et éradiquer des virus nouveaux. Dans notre prototype, des ordinateurs compatibles PC, connectés par un réseau à un ordinateur central, exécutent tous le programme *AntiVirus* d'*IBM*. Pour détecter la présence d'un virus, le programme *AntiVirus* utilise diverses méthodes heuristiques : il observe des modifications du comportement de l'ordinateur, recherche des modifications

suspectes des programmes ou analyse des signatures ; il utilise le réseau pour envoyer tout programme suspect à un ordinateur d'administration du réseau.

À la réception d'un programme potentiellement infecté, l'ordinateur d'administration l'envoie à un ordinateur d'analyse de virus qui se comporte comme une «boîte de Pétri» numérique : les programmes de cette machine de test piègent le virus à l'aide de programmes spécialement conçus, nommés «leurres», qui sont exécutés, copiés ou manipulés. Cette phase est particulièrement utile pour faire se découvrir les virus polymorphes. Ces virus sont des virus de macros dont les lignes de programme sont cryptées, de sorte qu'un logiciel antiviral classique est en général incapable de reconnaître sa signature. De plus, lorsqu'ils se dupliquent, les virus polymorphes utilisent un code de chiffrement généré aléatoirement : les signatures sont modifiées à chaque copie. Pour se propager, un virus doit infecter des programmes qui sont souvent utilisés, si bien que l'activité des leurres fait ressortir le code du virus. Ainsi le virus se découvre, et on peut l'éradiquer. Pendant cette phase, on peut aussi comprendre d'autres caractéristiques du virus.

Les leurres infectés sont analysés par d'autres éléments du système immunitaire, qui extraient la signature du virus et fournissent des directives pour le détecter et l'enlever. À partir d'un programme infecté, l'analyseur de virus met généralement moins de cinq minutes pour engendrer ses directives. La machine d'analyse envoie cette information à l'ordinateur qui a demandé de l'aide ; ce dernier incorpore l'information dans une base de données permanente contenant les remèdes à utiliser pour les virus connus. Il peut alors localiser et enlever toutes les copies du virus, et il est définitivement protégé de futures rencontres.

Lorsque l'ordinateur infecté est relié à d'autres machines par l'intermédiaire d'un réseau local, il est tout à fait possible que le virus les ait également envahis. Aussi les nouvelles directives de surveillance et d'éradication sont-elles envoyées automatiquement aux ordinateurs voisins, de sorte que chaque machine peut vérifier immédiatement la présence éventuelle du nouveau virus. Puisque les virus informatiques utilisent le réseau pour se répandre rapidement, l'utilisation du réseau fournit l'antidote aux machines qui en ont besoin. En envoyant aux clients d'un site infecté les dernières directives, on peut, en principe, immuniser très rapidement toute la population des ordinateurs contre l'émergence d'un nouveau virus.

Malgré les efforts, les virus informatiques cohabiteront avec nos ordinateurs. Chaque type de virus évoluera à sa façon, mais, collectivement, la coévolution des virus informatiques et des techniques antivirales se poursuivra comme, en biologie, le font les parasites et leurs hôtes. Ces deux composantes évolueront aussi en réaction aux changements de l'environnement informatique. Par exemple, les agents logiciels itinérants devront être protégés des ordinateurs qu'ils traversent, tout comme ces derniers de la malveillance de certains agents logiciels. Peut-être les virus et les systèmes immunitaires informatiques sont-ils les précurseurs d'un riche écosystème de vie artificielle où les proies et les prédateurs vivront, mourront et coopéreront.

Jeffrey KEPHART, Gregory SORKIN, David CHESS et Steve WHITE travaillent au Centre de recherche Thomas Watson de la Société *IBM*, à Yorktown Heights.

Rogue Programs : Viruses, Worms and Trojan Horses, sous la direction de Lance J. Hoffman, Van Nostrand Reinhold, 1990.

Frederick B. COHEN, *A Short Course on Computer Viruses*, deuxième édition, John Wiley & Sons, 1994.

Robert SLADE, *Robert Slade's Guide to Computer Viruses*, Springer-Verlag, 1994.

Jeffrey O. KEPHART, Gregory B. SORKIN, William C. ARNOLD, David M. CHESS, Gerald J. TESAURO et Steve R. WHITE, *Biologically Inspired Defenses against Computer Viruses*, in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montréal, 20-25 août 1995, Morgan Kaufmann Publishers, Inc.

Mark LUDWIG, *Du virus à l'antivirus : guide d'analyse*, Éditions Dunod, 1997.

Des liens vers de nombreux sites qui traitent des virus sont sur notre site Web : http://www.pourlascience.com

Le riz génétiquement modifié

PAMELA RONALD

Les biologistes utilisent le génie génétique pour protéger le riz contre les maladies.

Le riz est l'aliment le plus important du Globe : près de deux milliards de personnes ont une alimentation essentiellement à base de cette céréale. Les rizières couvrent plus de 144 millions d'hectares et procurent 560 millions de tonnes de grains par an. Toutefois on ne récolte pas tout le riz qui pousse : les insectes, les bactéries, les virus et les moisissures prélèvent une dîme notable.

La nielle bactérienne, notamment, est un des pires fléaux de l'Asie et de l'Afrique : cette maladie est due à des bactéries *Xanthomonas oryzae*, qui se propagent rapidement par l'intermédiaire de l'eau, de plant à plant, de champ à champ. Une fois infectées, les feuilles se couvrent de lésions, puis jaunissent et se flétrissent en quelques jours. Dans les champs le plus gravement touchés, la nielle détruit jusqu'à la moitié de la récolte.

Cependant les plants de riz possèdent parfois des gènes qui les protègent contre de nombreuses maladies, notamment contre la nielle. Les agriculteurs admettent qu'aucune variété n'a tous les gènes de résistance contre toutes les maladies ; certains plants sont plus résistants que d'autres à certaines maladies, moins à d'autres. Aussi, depuis près d'un siècle, les agronomes cherchent à conférer les performances des variétés rustiques aux variétés cultivées, mais la sélection par la génétique classique est lente, et les sélectionneurs mettent des décennies à transmettre des caractères d'une variété à une autre.

Grâce aux techniques du génie génétique, on introduit aujourd'hui des gènes de résistance directement dans des plants de riz, et l'on gagne ainsi des années de travail. Mes collègues et moi avons récemment cloné le premier de ces gènes de résistance du riz – un gène protégeant contre des formes classiques de nielle –, puis nous l'avons utilisé pour créer des plants résistants.

1. LES PLANTS DE RIZ sont sensibles à de nombreuses maladies, telle la nielle bactérienne, qui se manifeste par des lésions des feuilles et un rendement réduit *(à gauche)*. La maladie se propage par les gouttelettes d'eau contaminée par les bactéries, qui s'introduisent par des lésions des feuilles. En quelques jours, des lésions jaunes se développent sur les feuilles infectées. Les plants les plus jeunes succombent généralement à la maladie, sauf quand ils possèdent des gènes de résistance *(à droite)*.

Richard Jones