Vanderbilt Journal of Entertainment & Technology Law

Volume 10
Issue 2 *Issue 2 - Winter 2008*

Article 1

2008

Opinionated Software

Meiring de Villiers

VANDERBILT JOURNAL OF ENTERTAINMENT AND TECHNOLOGY LAW

VOLUME 10 WINTER 2008 NUMBER 2

Opinionated Software

*Meiring de Villiers**

ABSTRACT

Information security is an important and urgent priority in the computer systems of corporations, governments, and private users. Malevolent software, such as computer viruses and worms, constantly threatens the confidentiality, integrity, and availability of digital information. Virus detection software announces the presence of a virus in a program by issuing a virus alert. A virus alert presents two conflicting legal issues. A virus alert, as a statement on an issue of great public concern, merits protection under the First Amendment. The reputational interest of a plaintiff disparaged by a virus alert, on the other hand, merits protection under the law of defamation. The United States Supreme Court has struck a balance by constitutionalizing the common law of defamation in a series of influential decisions. This article focuses on two implications of these decisions, namely that (1) a plaintiff must show that the defamatory statement is objectively verifiable as true or false; and (2) a plaintiff

* John Landerer Faculty Fellow, University of New South Wales, Sydney, Australia; P.Eng, Canada. Ph.D., Economics, Stanford University, 1997; J.D., Stanford University, 1995; B.Sc., Electrical Engineering, University of Pretoria, South Africa. I would like to thank Professors Robert Rabin and George Winterton for their helpful comments.

must prove its falsity with convincing clarity, while the defendant may prove the truthfulness of the statement as a defense. The crucial issues in these implications are truth, falsity, and verifiability.

This article analyzes the balance between the conflicting legal rights associated with a virus alert. It focuses on the legal meanings of truth, falsity, and verifiability of a virus alert, and the resolution of these issues in the context of the technology involved in a virus alert. The analysis merges perspectives from constitutional law, the law of defamation, and information technology. Insights from theoretical computer science demonstrate, for instance, that the truth of a virus alert may be unverifiable. In such a case the alert would receive full constitutional protection under the Supreme Court's First Amendment defamation jurisprudence.

TABLE OF CONTENTS

Computer security is an important and urgent priority in the information networks of corporations, governments, and, increasingly, private users, especially in the aftermath of the terrorist attacks of September 11, 2001. The interconnectivity and interdependence of computers on the Internet have made users increasingly vulnerable to

cyber attacks emanating from a variety of wrongdoers, such as cyber criminals, terrorist groups, and, perhaps, even rogue nation states.[1]

The most powerful weapon available to cyber attackers is a type of computer code generically known as "malevolent software." Malevolent software is designed to disrupt the operation of computer systems.[2] The most common of these rogue programs are the computer virus, and its common variant, the so-called "worm."[3] Viruses can be programmed to access and steal confidential information; to corrupt and delete electronic data; and to monopolize computational resources that should be available to legitimate users.

The escalation of virus attacks on the Internet has prompted the development of advanced virus detection and elimination technologies. Virus detection software issues an alert when it detects virus-like behavior or properties in a program. The leading anti-virus technologies are sophisticated and effective, but virus detection errors, known as false positives and false negatives, nevertheless do occur. A false positive is an indication that a virus has been found when, in fact, there is none. A false negative is the converse, namely the failure to detect a virus when one is actually present.

A virus alert tends to harm the reputation of a corporation whose software product has been tagged as viral. The maligned corporation may initiate a defamation action against the

---

1. *See, e.g., Overview of the Cyber Problem — A Nation Dependent and Dealing with Risk: Hearing of the Subcomm. on Cybersecurity, Science, and Research and Development Before the H. Select Comm. on Homeland Security*, 108th Cong. 22 (2003) (statement of Richard D. Pethia, Director, CERT® Centers, Software Engineering Institute, Carnegie Mellon University) ("As critical infrastructure operators strive to improve their efficiency and lower costs, they are connecting formerly isolated systems to the Internet to facilitate remote maintenance functions and improve coordination across distributed systems. Operations of the critical infrastructures are becoming increasingly dependent on the Internet and are vulnerable to Internet based attacks."); *see also* DOROTHY E. DENNING, INFORMATION WARFARE *AND* SECURITY 17 (1999) ("Through increased automation and connectivity, the critical infrastructures of a country become increasingly interdependent. Computers and telecommunications systems, for example, support energy distribution, emergency services, transportation, and financial services.").

2. ROBERT SLADE, DICTIONARY OF INFORMATION SECURITY 118 (2006).

3. A computer virus can be described as a program that (i) infects a host program by attaching itself to the host, (ii) executes when the host is executed, and (iii) spreads by cloning itself and attaching the clones to other host programs. Viruses often also have a so-called "payload," capable of harmful side-effects, such as deleting, stealing, or modifying information. *See* FREDERICK B. COHEN, A SHORT COURSE ON COMPUTER VIRUSES 1-2 (2d ed. 1994); DOROTHY E. DENNING & PETER J. DENNING, INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS 73-75 (1998). A "worm" is similar to a virus in most respects, except that it does not need to attach itself to a host program to replicate and spread. Like viruses, worms may carry destructive payloads. *See generally* John F. Schoch & Jon A. Hupp, *The "Worm" Programs—Early Experience with a Distributed Computation*, 25 COMM. ACM 172, 172 (1982).

manufacturer of the virus detection software. For example, in December 1992, a news story in the *Wall Street Journal* reported that a federal judge had ordered McAfee Associates, Inc., a producer of computer security software, to stop distribution of one of its products that falsely identified a virus in the software of a company, Imageline, Inc.[4] Imageline sued McAfee, alleging defamation, among other claims.[5] The complaint alleged that the false positives scared customers away, hurting the company's reputation.[6] McAfee declined comment, other than stating that the suit was without merit.[7]

This article analyzes the balance between two conflicting legal rights associated with a virus alert. A virus alert, as a statement on an issue of great public concern, merits protection under the First Amendment. The reputational interest of a plaintiff disparaged by a virus alert, on the other hand, merits protection under the law of defamation. The United States Supreme Court has struck a balance by constitutionalizing the common law of defamation in a series of decisions, starting in 1964 with *New York Times Co. v. Sullivan*.[8] This article focuses on two implications of these decisions, namely that (1) a plaintiff must show that the defamatory statement is objectively verifiable as true or false; and (2) a plaintiff must prove the statement's falsity with convincing clarity, while the defendant may prove the truthfulness of the statement as a defense. The crucial issues in these implications are truth, falsity, and verifiability.

The analysis merges three perspectives, namely the Supreme Court's First Amendment defamation jurisprudence, the common law of defamation, and the technological environment in which a virus alert occurs. Analysis of the defamatory implication of a virus alert shows that a virus alert is substantially true if, and only if, the object identified as viral is capable of reproducing by executing an infection module. Conversely, a virus alert is false if the object either does not have an infection module or if the infection module cannot execute, perhaps due to a programming or logical error. This result provides a rigorous and logical definition of the truthfulness of a virus alert as a

---

4. Junda Woo, *False Alarms Over a Virus*, WALL ST. J., Dec. 29, 1992, at B6.

5. John Burgess, *Viruses: An Overblown Epidemic?; Suit Against a Calif. Firm Highlights Computer Industry Debate*, WASH. POST, Dec. 30, 1992, at F1; Kephart et al., *Blueprint for a Computer Immune System*, IBM THOMAS J. WATSON RESEARCH CENTER 1, 11 (1997), *available at* http://www.research.ibm.com/antivirus/SciPapers/Kephart/VB97/index.html.

6. *Id.*

7. *Id.*; *see* Woo, *supra* note 4, at B6.

8. 376 U.S. 254 (1964).

defamatory statement. It also provides the forensic basis for proof of truthfulness.

This article further demonstrates the resolution of the issues of truth, falsity, and verifiability based on forensic analysis of (1) the technology that issued a virus alert and (2) the digital properties of the viral object. The analysis highlights a striking entanglement of law and technology. Insights from theoretical computer science show, for instance, that the truth of a virus alert, as defined in this article, may be indeterminate under certain conditions. When these conditions apply, the alert would receive full protection under the Supreme Court's First Amendment defamation jurisprudence.

This article is organized as follows: Part I discusses the elements of a defamation action; Part II reviews the evolution of the Supreme Court's First Amendment defamation jurisprudence; Part III discusses the principles of malevolent software; Part IV analyzes the anatomy of a virus alert; and Part V analyzes the truth, falsity, and verifiability of a virus alert.

I. PRINCIPLES OF DEFAMATION

The tort of defamation protects the interest of a person or corporation in their reputation and good name.[9] A defamatory statement is a false statement of fact about a person or business entity that tends to harm their reputation, respect, or goodwill.[10] Courts have upheld claims for defamation when a party has untruthfully stated that a person is a credit risk,[11] that a kosher meat dealer has

9. RESTATEMENT (SECOND) OF TORTS § 559 (1977). Defamation is the broader term for libel and slander. Libel is concerned with written or printed words, or more generally, embodiment of the defamatory message in tangible or permanent form. W. PAGE KEETON ET AL., PROSSER & KEETON ON THE LAW OF TORTS § 112, at 785 (5th ed. 1984). Slander constitutes oral defamation. *Id.*

10. RESTATEMENT (SECOND) OF TORTS §§ 558-559 ("A communication is defamatory if it tends so to harm the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him."); Jessica R. Friedman, *Defamation*, 64 FORDHAM L. REV. 794 (1995). The Code of the Australian state of Queensland defines "defamation" as:

> Any imputation concerning any person, or any member of the person's family, whether living or dead, by which the reputation of that person is likely to be injured, or by which that person is likely to be injured in the person's profession or trade, or by which other persons are likely to be induced to shun or avoid or ridicule or despise the person . . . .

Queensl. Stat., c. 35, § 366 (1995).

11. *See, e.g.*, Neaton v. Lewis Apparel Stores, Inc., 48 N.Y.S.2d 492, 497 (N.Y. App. Div. 1944) (reversing the lower court's dismissal of a claim for libel based on a letter written by defendant to plaintiff's employer stating that plaintiff was a bad credit risk).

sold bacon,[12] and that a physician has advertised.[13] Defamation law aims to protect the reputational interests of a plaintiff by allowing her to restore her good name, and to obtain compensation and redress for harm caused by defamatory statements.[14] Courts have extended the protection of defamation law to the reputational interests of corporations.[15] Although a corporation has no reputation in the personal sense of an individual,[16] it has a reputation and standing in the business community in which it operates.[17] A corporation can sue for defamatory statements related to matters affecting its business reputation and practices, such as financial soundness, management, and efficiency.[18]

The complexity of the tort of defamation is illustrated by the elements that have to be satisfied to establish a cause of action. One author has identified nine elements,[19] while another lists twenty-three,[20] each crucial to a defamation action. The defamation plaintiff must plead and prove the following elements:[21]

1. The statement of fact must be published to a third party other than the plaintiff.

2. The statement must be false.

---

12. *See* Braun v. Armour & Co., 173 N.E. 845, 845 (N.Y. 1930).

13. *See* Gershwin v. Ethical Publ'g Co., 1 N.Y.S.2d 904, 906 (N.Y. City Ct. 1937).

14. Milkovich v. Lorain Journal Co., 497 U.S. 1, 12 (1990) ("Defamation law developed not only as a means of allowing an individual to vindicate his good name, but also for the purpose of obtaining redress for harm caused by such statements." (citation omitted)).

15. *See* RESTATEMENT (SECOND) OF TORTS § 561 cmt. b (1977); KEETON ET AL., *supra* note 9, § 111, at 779; *see also* Brown & Williamson Tobacco Corp. v. Jacobson, 827 F.2d 1119 (7th Cir. 1987).

16. *See* Golden Palace, Inc. v. Nat'l Broad. Co., 386 F. Supp. 107, 109 (D.D.C. 1974); Di Giorgio Fruit Corp. v. AFL-CIO, 30 Cal. Rptr. 350, 355-56 (Cal. Dist. Ct. App. 1963); Reporters' Ass'n of Am. v. Sun Printing & Publ'g Ass'n, 79 N.E. 710, 711 (N.Y. 1906).

17. *See Di Giorgio Fruit Corp.*, 30 Cal. Rptr. at 356.

18. *See* Diplomat Elec., Inc. v. Westinghouse Elec. Supply Co., 378 F.2d 377, 382-83 (5th Cir. 1967); Aetna Life Ins. Co. v. Mut. Benefit Health & Accident Ass'n, 82 F.2d 115, 119 (8th Cir. 1936); Maytag Co. v. Meadows Mfg. Co., 45 F.2d 299, 302 (7th Cir. 1930); *Di Giorgio Fruit Corp.*, 30 Cal. Rptr. at 355-56; RESTATEMENT (SECOND) OF TORTS § 561 (1977); KEETON ET AL., *supra* note 9, § 128, at 962-63 (discussing the law of injurious falsehoods' concern with false statements that harm economic interests, but do not harm the corporate reputation); Milo Geyelin, *Corporate Mudslinging Gets Expensive—Aggrieved Do More Than Turn Other Cheek*, WALL ST. J., Aug. 4, 1989, at B1.

19. *See* Robert D. Nelon, *Media Defamation in Oklahoma: A Modest Proposal and New Perspectives—Part I*, 34 OKLA. L. REV. 478, 487-88 (1981).

20. W. Page Keeton, *Defamation and Freedom of the Press*, 54 TEX. L. REV. 1221, 1233-35 (1976).

21. *See* 1 RODNEY A. SMOLLA, LAW OF DEFAMATION § 1:34 (2d ed. 1986 & Supp. 2007); Friedman, *supra* note 10, at 794.

3. The statement must be defamatory or, in other words, harmful to the reputation of the plaintiff.

4. The statement must have reasonably referred to the plaintiff.

5. The defendant must have acted with the requisite degree of fault. The fault requirement depends on the plaintiff's status. A private-figure plaintiff must prove negligence, namely that, by a preponderance of the evidence, the defendant lacked reasonable grounds for believing the statement to be true, or failed to take reasonable care to ascertain the truth. A public plaintiff[22] must prove by clear and convincing evidence that the defendant published the statement with "actual malice." Actual malice is defined as "with knowledge that the statement was false, or with reckless disregard of whether it was false or not."[23]

6. The statement must be objectively capable of being proven materially false.

7. The statement must have caused actual harm to the plaintiff. There are three categories of defamation damages: special, presumed, and punitive damages. Special damages compensate the plaintiff for pecuniary or economic loss flowing directly from the reputational harm caused by the defamatory statement.[24] This type of harm must be proven with reasonable certainty.[25] In cases where damages are difficult to quantify, a plaintiff may be allowed to recover presumed damages, if certain conditions are met.[26] The plaintiff may recover punitive damages in cases where the defendant published a statement with knowledge of its falsity or with reckless disregard for its truth or falsity.[27]

8. The statement must not be privileged, as a privileged publication is not actionable. Judges, attorneys, jurors, and legislators, for instance, can plead an absolute privilege for statements made in furtherance of their official duties.[28]

A person or corporation can be defamed by more than written or spoken words.[29] Defamation may occur by means of a picture, a

22. "Public plaintiff" includes a public official, see *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964); a public figure, see *Curtis Publishing Co. v. Butts*, 388 U.S. 130 (1967); and a limited purpose public figure, see *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 345 (1974). Limited purpose public figures are people who "have thrust themselves to the forefront of particular public controversies in order to influence the resolution of the issues involved." *Gertz*, 418 U.S. at 345. The "public official" category is fairly wide, and includes, for instance, government employees. *See* 1 SLADE R. METCALF & LEONARD M. NIEHOFF, RIGHTS AND RESPONSIBILITIES OF PUBLISHERS, BROADCASTERS, AND REPORTERS 1.50, at 177 (2002 & Supp. 2006).

23. *New York Times*, 376 U.S. at 279-80.

24. 1 ROBERT D. SACK, SACK ON DEFAMATION: LIBEL, SLANDER AND RELATED PROBLEMS § 10.3.2 (3d ed. 2007).

25. *See* Matherson v. Marchello, 473 N.Y.S.2d 998, 1000 (N.Y. App. Div. 1984).

26. Presumed damages may be allowed, even if special damages cannot be proven, provided the defamation falls into a "per se" category. A statement that the plaintiff had committed a crime, for instance, would be defamation per se. *See* RESTATEMENT (SECOND) OF TORTS § 570 (1977).

27. KEETON ET AL., *supra* note 9, § 115, at 845.

28. *Id.* § 115, at 824-32.

29. The Restatement supports a broad interpretation of what may constitute defamatory speech. *See* RESTATEMENT (SECOND) OF TORTS § 565 cmt. b, at 170 ("To be

gesture, a loaded question, or an insinuation. And, the defamatory imputation may be indirect. For example, signing the plaintiff's name to false[30] or bad authorship[31] has been held to be defamatory. Plaintiffs who were defamed by comments published on the Internet have filed successful defamation actions. For instance, in *Stratton Oakmont, Inc. v. Prodigy Services Co.*, a defamatory statement was posted on a bulletin board maintained by Prodigy, an Internet service provider (ISP).[32] The statement claimed that the plaintiffs had committed fraud in connection with an initial public stock offering.[33] The plaintiffs filed a defamation action, and prevailed in the Supreme Court of New York.[34] Congress subsequently passed legislation exempting ISPs from liability for online defamation.[35] Shortly thereafter, in *Blumenthal v. Drudge*, the United States District Court for the District of Columbia denied recovery to a plaintiff in a defamation action against an ISP under the safe harbor provision for ISPs.[36] The court observed that, although an ISP is immune from liability, the original author of the defamatory statements could potentially be held liable.[37]

---

defamatory under the rule stated in this Section, it is not necessary that the accusation or other statement be by words. It is enough that the communication is reasonably capable of being understood as charging something defamatory."); *see also* Defamation Act, 1996, c. 31, § 17(1) (Eng.) (stating that a defamatory statement means "words, pictures, visual images, gestures or any other method signifying meaning").

30. *See, e.g.*, Ben-Oliel v. Press Publ'g Co., 167 N.E. 432, 434 (N.Y. 1929) ("To publish in the name of a well-known author any literary work, the authorship of which would tend to injure an author holding his position in the world of letters, has been held to be a libel."); Locke v. Benton & Bowles, Inc., 1 N.Y.S.2d 240 (N.Y. Sup. Ct. 1937) (holding that falsely attributing authorship of a script states a valid claim for defamation), *rev'd on other grounds*, 2 N.Y.S.2d 150, 151-152 (N.Y. App. Div. 1938) (dismissing for failure to set forth appropriate facts in the complaint).

31. *See, e.g.*, Sperry Rand Corp. v. Hill, 356 F.2d 181 (1st Cir. 1966) (vacating judgment for plaintiff on libel and invasion of privacy claims and remanding to determine compensatory damages for libel based on false attribution of authorship of an article); Carroll v. Paramount Pictures, 3 F.R.D. 47 (S.D.N.Y. 1943) (denying defendant–movie studio's motion for summary judgment in action for libel claiming that defendant falsely attributed production of a movie to plaintiff–producer).

32. Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 805178, at *1 (N.Y. Sup. Ct. Dec. 11, 1995).

33. *Id.*

34. *Id.*

35. Communications Decency Act of 1996, Pub. L. No. 104-104, 110 Stat. 56 (codified as amended at 47 U.S.C. § 230 (2000)).

36. 992 F. Supp. 44, 50 (D.D.C. 1998); *see also* Zeran v. Am. Online, Inc., 129 F.3d 327 (4th Cir. 1997) (immunizing ISPs to defamation liability for third-party content).

37. *Blumenthal*, 992 F. Supp. at 51 ("None of this means, of course, that the original culpable party who posts defamatory messages would escape accountability." (quoting *Zeran*, 129 F.3d at 331)).

The defamatory comments in Internet defamation cases were published as computer-generated words and images, which were downloaded and displayed on user terminals. A virus alert is communicated to a recipient in a similar format, namely a computer-generated message alerting a computer user to the presence of a virus. The analogy suggests that courts will likely recognize that a plaintiff can be defamed by a virus alert.

The defendant in a defamation action involving a virus alert may argue that computer-generated communication, such as a virus alert, merits protection under the First Amendment. This position has received support among academic commentators,[38] and courts have in fact recognized protection for specific categories of computer-generated output, such as digital simulation of sexual activity by minors that does not rise to the level of obscenity.[39]

Professor Dan Burk has argued in favor of First Amendment protection for computer-generated output, pointing to the analogy between computer output and First Amendment protected music.[40] The scope of First Amendment protection of music extends to the musical output of a piano roll or compact disc. The output generated by computers, as the digital analogue of the output of a piano roll, whether in the form of text, graphics, or sound, should, therefore, receive equivalent protection.[41] Professor Burk explains the analogy by noting that piano rolls have sequences and patterns of punched

38. Academic commentators have argued that computer output is an expression of functions and operations performed by a computer, analogous to spoken and written expressions of the human mind, and thus, within the scope of First Amendment protection. *See* Roy N. Freed, *Products Liability in the Computer Age*, 17 JURIMETRICS J. 270, 280 (1976-77); *see also* Gary T. Walker, *The Expanding Applicability of Strict Liability Principles: How is a "Product" Defined?*, 22 TORT TRIAL & INS. PRAC. L.J. 1, 12-15 (1986).

39. *See* Ashcroft v. Free Speech Coal., 535 U.S. 234 (2002) (finding a federal statute that banned computer-generated child pornography to be unconstitutional); Norman T. Deutsch, *Professor Nimmer Meets Professor Schauer (and Others): An Analysis of "Definitional Balancing" as a Methodology for Determining the "Visible Boundaries of the First Amendment,"* 39 AKRON L. REV. 483, 524 (2006) ("'[T]he distribution of descriptions or other depictions of sexual conduct [by minors], not otherwise obscene, which do not involve live performance or other visual reproduction of live performances, retains First Amendment protection.' *This includes . . . computer generated images . . . .*" (quoting New York v. Ferber, 458 U.S. 747, 764-65 (1982) (emphasis added) (alteration in original))); Norman Andrew Crain, Commentary, *Bernstein, Karn and Junger: Constitutional Challenges to Cryptographic Regulations*, 50 ALA. L. REV. 869, 887 (1999) ("[E]xpression does not lose First Amendment protection just because it interacts with a machine or . . . a computer."). However, obscene works, including computer-generated images involving obscenity, are not First Amendment protected. *See Free Speech Coal.*, 535 U.S. at 240.

40. *See* Ward v. Rock Against Racism, 491 U.S. 781, 790 (1989) ("Music, as a form of expression and communication, is protected under the First Amendment.").

41. Dan L. Burk, *Patenting Speech*, 79 TEX. L. REV. 99, 115 (2000).

holes that "constitute a type of machine-readable 'program'" and "express music by tripping the mechanism of a player piano."[42]

However, courts have been reluctant to extend constitutional protection to computer-generated output that does not advance the ideals of the First Amendment. In *Commodity Futures Trading Commission v. Vartuli*,[43] the Second Circuit contemplated whether the First Amendment protected computer-generated trading commands. The computer system at issue required no independent intellectual effort from the user.[44] The user, for the system to work as marketed, was supposed to obey the computer's buy and sell signals literally and without question.[45] The court stated that the purpose of the computer output was not to communicate information, but to prompt action without engaging the mind or will of the recipient.[46] None of the ideals pursuant to which speech is normally accorded constitutional protection—such as the pursuit of truth, the prevention of abuse of authority, and the functioning of a democracy—were relevant to this communication.[47] The court concluded that the defendant who distributed this automatic trading system did not engage in constitutionally protected speech.[48]

In conclusion, a virus alert, as a statement on a significant public issue, merits First Amendment protection, but is also subject to the law of defamation.

II. EVOLUTION OF FIRST AMENDMENT DEFAMATION JURISPRUDENCE

Until 1964, defamation was outside the scope of First Amendment protection. Defamation law strongly favored the plaintiff, and courts treated defamation virtually as a strict liability tort.[49] A

---

42. *Id.*

43. 228 F.3d 94 (2d Cir. 2000).

44. *Id.* at 111.

45. *Id.*

46. *Id.*

47. *Id.*

48. *Id.* The court issued a caveat:

> Statements in the form of orders or instructions are strikingly common. . . . We do not think and do not mean to suggest by our holding today that such communications "can claim talismanic immunity from constitutional limitations." . . . Any assertion that a statement like or unlike the "buy" or "sell" instructions issued by a . . . computer is not fully protected by the Constitution should be subjected to careful and particularized analysis to insure that no speech entitled to First Amendment protection fails to receive it.

Id. at 112 (quoting N.Y. Times Co. v. Sullivan, 376 U.S. 254, 269 (1963)).

49. *See* Lackland H. Bloom, Jr., *Proof of Fault in Media Defamation Litigation*, 38 VAND. L. REV. 247, 249 (1985); Marc A. Franklin & Daniel J. Bussel, *The Plaintiff's Burden in Defamation: Awareness and Falsity*, 25 WM. & MARY L. REV. 825, 826 (1984).

plaintiff merely had to allege falsity to establish a cause of action for defamation, while defendants had to prove the truth.[50] Liability and damages were presumed, and a plaintiff could recover without showing any actual harm.[51] Defendants had several potential defenses, including truth, absolute privilege, conditional privilege, and fair comment.[52] In practice, however, these defenses were difficult to establish, and pleading them sometimes exposed defendants to further liability.[53]

The constitutionalization of the common law of defamation, which started in 1964, dramatically reshaped plaintiffs' positions, especially with respect to the burdens of proof of truth and falsity, fault, and the opinion privilege.

A. Truth, Falsity, and the Burden of Proof

In a landmark decision in *New York Times Co. v. Sullivan*,[54] the Supreme Court redefined the contours of libel litigation and eroded much of plaintiffs' previously favored positions. In *New York Times*, a public official of Alabama filed a defamation suit against the New York Times based on an advertisement in the newspaper that alleged police misconduct towards members of the civil rights movement.[55] The trial court found for the plaintiff and awarded damages of $500,000.[56] The Supreme Court of Alabama affirmed the judgment.[57]

The United States Supreme Court reversed, holding that, in defamation actions brought by public officials, the Constitution requires a plaintiff to show by clear and convincing evidence that the

---

50. *See* Bloom, *supra* note 49, at 249; Franklin & Bussel, *supra* note 49, at 826.

51. *See, e.g.*, Lewis v. Hayes, 171 P. 293, 294 (Cal. 1918).

52. *See* Rodney W. Ott, Note, *Fact and Opinion in Defamation: Recognizing the Formative Power of Context*, 58 FORDHAM L. REV. 761, 763 (1990) ("A defendant could invoke a fair comment privilege by proving that (1) the statement concerned a matter of legitimate public interest, (2) the facts upon which the statement was based were either stated or known to the reader, (3) the statement was the actual opinion of the defendant, and (4) the statement was not motivated solely by the purpose of causing harm to the plaintiff.").

53. Franklin & Bussel, *supra* note 49, at 826 n.6 ("In addition to the difficulty with regard to proof, an assertion in the pleadings that the statement was true may expose the defendant to further liability. If he should fail to prevail on that issue, the court may consider the pleading to be a republication of the libel." (citation omitted)).

54. 376 U.S. 254 (1964).

55. *Id.* at 256.

56. *Id.*

57. New York Times Co. v. Sullivan, 144 So. 2d 25, 52 (Ala. 1962), *rev'd*, 376 U.S. 254 (1964).

statement at issue was published with "actual malice,"[58] a standard that the plaintiff had not met.[59] The Court defined actual malice as a statement made with either knowledge of its falsity or reckless disregard for the truth.[60] The Court further held that a plaintiff must prove actual malice by clear and convincing evidence, which is a stricter standard than the civil preponderance of the evidence, but less rigorous than the criminal standard of beyond a reasonable doubt.[61] The rationale underlying *New York Times* was that the First Amendment should function "'to assure unfettered interchange of ideas for the bringing about of political and social changes desired by the people.'"[62] The Court concluded that imposing the burden of proving the truth of the statement on the critic of official conduct amounted to a form of self-censorship, contrary to the ideals of the First Amendment.[63]

Although the *New York Times* Court did not specifically state that truth is an absolute defense in a defamation action against a public official, it is implied by the actual malice requirement.[64] The logic of this conclusion is well articulated in *Rinaldi v. Holt, Rinehart & Winston, Inc.*, where the Court of Appeals of New York reasoned that placing the burden of proof of falsity on the plaintiff "follows naturally from the actual malice standard. Before knowing falsity or reckless disregard for truth can be established, the plaintiff must establish that the statement was, in fact, false."[65] The *New York Times* decision has also been interpreted by the Supreme Court,[66]

---

58. New York Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).

59. *Id.* at 283.

60. *Id.* at 280. In a subsequent opinion, the Court described "reckless disregard" for the truth as entertaining serious doubts about the truth of the statement before making it. St. Amant v. Thompson, 390 U.S. 727, 731 (1968).

61. *New York Times*, 376 U.S. at 279-80.

62. *Id.* at 269 (quoting Roth v. United States, 354 U.S. 476, 484 (1957)).

63. *See id.* at 279.

64. *See* 1 SMOLLA, *supra* note 21, § 5:4.

65. 366 N.E.2d 1299, 1306 (N.Y. 1977).

66. *See* Herbert v. Lando, 441 U.S. 153, 176 (1979) ("The plaintiff's burden is now considerably expanded. In every or almost every case, the plaintiff must focus on the editorial process and prove a false publication attended by some degree of culpability on the part of the publisher."); Cox Broad. Corp. v. Cohn, 420 U.S. 469, 490 (1975) ("[T]he defamed public official or public figure must prove not only that the publication is false but that it was knowingly so or was circulated with reckless disregard for its truth or falsity."); Garrison v. Louisiana, 379 U.S. 64, 74 (1964) ("We held in *New York Times* that a public official might be allowed the civil remedy only if he establishes that the utterance was false . . . or in reckless disregard of whether it was false or true."). The *New York Times* Court itself clearly stated that true speech can never be the basis of liability. *New York Times*, 376 U.S. at 271 ("Authoritative interpretations of the First Amendment guarantees have

lower courts,[67] and academic commentators[68] as imposing on a public figure plaintiff the burden of proof of falsity. Public figures are people who are in the public eye, but not public officials.[69]

In the same year that *New York Times* was decided, the Supreme Court considered the constitutionality of a Louisiana statute that allowed truth as a defense only for statements made "with good motives and for justifiable ends."[70] This limitation on the truth defense appears to have been held unconstitutional when the Court declared that "[t]ruth may not be the subject of either civil or criminal sanctions where discussion of public affairs is concerned."[71] Three years later, in *Curtis Publishing Co. v. Butts*,[72] the Court extended the actual malice standard of *New York Times* to public figure defamation plaintiffs. The Court also held that a public figure plaintiff must prove malice by clear and convincing evidence.[73]

In *Gertz v. Robert Welch, Inc.*, the plaintiff, Elmer Gertz, filed a defamation suit against a magazine that had made several untrue statements about him, including a charge that he was an official in a Communist organization that advocated the violent overthrow of the U.S. government.[74] Gertz prevailed at trial and won a jury award.[75] The trial court overturned the jury verdict, however, holding that the *New York Times* fault standard of actual malice applied to defamation actions involving matters of public concern, a standard that Gertz had not met.[76] The Supreme Court disagreed, reasoning that, although

---

consistently refused to recognize an exception for any test of truth . . . and especially one that puts the burden of proving truth on the speaker." (citation omitted)).

67. *See, e.g.*, Goldwater v. Ginzburg, 414 F.2d 324, 338 (2d Cir. 1969) (stating that "when the suit is brought by a public official or by a public figure . . . the burden of establishing that the published material was false is on the plaintiff"); Beckham v. Sun News, 344 S.E.2d 603, 604 (S.C. 1986) ("When a libel action is brought by a public official or public figure, the constitutional guarantees of freedom of speech and press require the plaintiff to establish the defamatory falsehood was made with actual malice, i.e., with knowledge of falsity or reckless disregard of whether it was false or not." (citation omitted)).

68. *See* Franklin & Bussel, *supra* note 49, at 851-54; Kathryn Dix Sowle, *Defamation and the First Amendment: The Case for a Constitutional Privilege of Fair Report*, 54 N.Y.U. L. REV. 469, 488 (1979); Linda Kalm, Note, *The Burden of Proving Truth or Falsity in Defamation: Setting a Standard for Cases Involving Nonmedia Defendants*, 62 N.Y.U. L. REV. 812, 813 (1987).

69. *See* Curtis Publ'g Co. v. Butts, 388 U.S. 130, 155 (1967).

70. *Garrison*, 379 U.S. at 70.

71. *Id.* at 74.

72. 388 U.S. 130 (1967).

73. *See id.* at 164.

74. 418 U.S. 323, 326 (1974).

75. *Id.* at 329.

76. *Id.*

the matter under litigation was one of public concern, Gertz was nevertheless a private figure because he had not deliberately brought himself into the public eye.[77] The Court found that a private figure's reputational interest merited greater protection than that provided by the actual malice requirement.[78] The Court concluded that states may impose any standard of care (other than strict liability) in defamation actions involving private plaintiffs.[79] Therefore, a private individual must prove that the defendant acted at least negligently.

In *Time, Inc. v. Firestone*, the Supreme Court indicated that truth would be a complete constitutional defense, in both private and public figure cases.[80] This position was consistent with the *Restatement (Second) of Torts*,[81] as well as a number of lower court holdings.[82] In *Philadelphia Newspapers, Inc. v. Hepps*, the Supreme Court confirmed the status of truth as a constitutional defense in private figure cases, at least where speech on matters of public interest is concerned and a media defendant is involved.[83]

B. Evolution of the Opinion Privilege

The general consensus is that expression of opinion, as distinct from statement of fact, must be protected from liability under defamation law.[84] Professor Robert Post commented that "opinions are in their nature debatable. To impose sanctions for 'false' opinions is to use the force of law to end this potential debate by imposing legally definitive interpretations of the cultural standards at issue."[85] Although support for an opinion privilege is evidently strong, the legal

---

77. *See id.* at 345.

78. *Id.* at 343-46.

79. *Id.* at 347.

80. 424 U.S. 448, 455 (1976).

81. RESTATEMENT (SECOND) OF TORTS § 581A (1977) ("One who publishes a defamatory statement of fact is not subject to liability for defamation if the statement is true.").

82. *See, e.g.*, Corabi v. Curtis Publ'g Co., 273 A.2d 899, 908 (Pa. 1971) ("[T]he opposite of falsity, truth, is a complete and absolute defense to a civil action for libel." (citations omitted)); *see also* 1 SMOLLA, *supra* note 21, § 5:5-5:8 (discussing truth as a constitutional defense in private figure cases).

83. 475 U.S. 767, 777-78 (1986).

84. *See* Bose Corp. v. Consumers Union of U.S., Inc., 466 U.S. 485, 503-04 (1984) ("The First Amendment presupposes that the freedom to speak one's mind is not only an aspect of individual liberty—and thus a good unto itself—but also is essential to the common quest for truth and the vitality of society as a whole.").

85. *See* Robert C. Post, *The Constitutional Concept of Public Discourse: Outrageous Opinion, Democratic Deliberation, and* Hustler Magazine v. Falwell, 103 HARV. L. REV. 603, 664 (1990).

distinction between actionable fact and protected opinion has proved to be difficult and elusive.[86]

Historically, treatment of opinion in the law of defamation has gone through three stages: (1) common law "fair comment"—largely prior to 1974; (2) protection based on the Supreme Court's dictum in *Gertz v. Robert Welch, Inc.*[87]—between 1974 and 1990; and (3) treatment based on the Court's opinion in *Milkovich v. Lorain Journal Co.*—1990 to present.[88]

The privilege of "fair comment" was born of the Court's sensitivity to the dangers inherent in legal limitations on freedom of expression. The privilege was designed to insure robust and open debate on public issues.[89] The defendant could rely on a fair comment privilege, provided the statement was (1) on a matter of public interest; (2) true or privileged; (3) the actual opinion of the speaker; and (4) made in good faith.[90] The privilege turned out to be inadequate and impractical, and its scope was uncertain.[91] For instance, a prediction as to whether a given statement merited protection would depend on factors that vary among jurisdictions.

The doctrine of fair comment was eventually superseded when the Supreme Court, in *Gertz v. Robert Welch, Inc.*, hinted that the traditional common law distinction between fact and opinion may also trigger First Amendment concerns.[92] Justice Powell, writing for the majority, elaborated on the fact-opinion distinction:

> Under the First Amendment there is no such thing as a false idea. However pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries but on the competition of other ideas. But there is no constitutional value in false statements of fact. Neither the intentional lie nor the

86. *See* Ott, *supra* note 52, at 761 n.1 (providing a bibliography on the pre–1990 literature on the fact-opinion distinction).

87. 418 U.S. 323 (1974).

88. 497 U.S. 1 (1990); *see* 1 SACK, *supra* note 24, § 4.2.1.

89. *See* Alfred Hill, *Defamation and Privacy under the First Amendment*, 76 COLUM. L. REV. 1205, 1227-36 (1976), *cited in* 1 SACK, *supra* note 24, § 4.2.2 n.18.

90. *See, e.g.*, Salinger v. Cowles, 191 N.W. 167, 173-74 (Iowa 1923) ("'The utmost extent to which these cases go is that, where a person, knowing or honestly believing that a candidate for public office is guilty of conduct affecting his fitness for the position to which he aspires, communicates that knowledge or belief to the electors whose support the candidate seeks, acting in good faith in the discharge of his duty to the public, the communication is privileged—a doctrine the correctness of which we need not now consider.'" (quoting Morse v. Times-Republican Printing Co., 100 N.W. 867, 873 (Iowa 1904))).

91. *See* W. Andrew Scott, Note, *Fair Comment in California: An Unwelcome Guest*, 57 S. CAL L. REV. 173, 195 (1983).

92. 418 U.S. 323.

> careless error materially advances society's interest in "uninhibited, robust, and wide-open" debate on public issues.[93]

Although technically dictum, Justice Powell's statement rapidly assumed constitutional status in the judiciary. Subsequent Supreme Court opinions have mentioned the *Gertz* dictum with approval,[94] and most state and federal courts have taken it to establish an absolute constitutional privilege for statements of opinion.[95] A subsequent version of the *Restatement (Second) of Torts* stated that "[t]he common law rule that an expression of opinion of the . . . pure[] type may be the basis of an action for defamation now appears to have been rendered unconstitutional by U.S. Supreme Court decisions."[96]

The *Gertz* dictum did not provide any analytical means of distinguishing between an actionable assertion of fact and protected opinion, and post-*Gertz* courts struggled with the distinction. In one influential decision, *Ollman v. Evans*, the District of Columbia Circuit formulated a widely used test.[97] Writing for the court, then-Judge Kenneth Starr articulated four factors that distinguished fact from opinion, namely (1) the ordinary meaning of the language used; (2) the verifiability of the statement; (3) its linguistic content; and (4) the social context.[98]

In *Milkovich v. Lorain Journal Co.*, the Supreme Court revisited the opinion privilege.[99] The plaintiff in Milkovich was a high school wrestling coach whose team had become involved in an altercation during a wrestling match.[100] The Ohio High School Athletic Association (OHSAA) conducted a hearing into the incident, in which Milkovich, as well as H. Don Scott, the Superintendent of

---

93. *Id.* at 339-40 (citation omitted).

94. *See* Hustler Magazine v. Falwell, 485 U.S. 46, 51 (1988); Bose Corp. v. Consumers Union of U.S., Inc., 466 U.S. 485, 504 (1984).

95. *See* Potomac Valve & Fitting, Inc. v. Crawford Fitting Co., 829 F.2d 1280, 1286 (4th Cir. 1987) ("The constitutional distinction between fact and opinion is now firmly established in the case law of the circuits."); Ollman v. Evans, 750 F.2d 970, 975 (D.C. Cir. 1984) ("*Gertz*'s implicit command thus imposes upon both state and federal courts the duty as a matter of constitutional adjudication to distinguish facts from opinions in order to provide opinions with the requisite, absolute First Amendment protection."); 1 SACK, *supra* note 24, § 4.2.3.1 ("By 1990 every federal circuit and the courts of at least thirty-six states and the District of Columbia had held that opinion is constitutionally protected because, according to *Gertz*, '[u]nder the First Amendment there is no such thing as a false idea.'" (alteration in original)).

96. RESTATEMENT (SECOND) OF TORTS § 566 cmt. c (1977). The Restatement defined "pure opinions" as those that "do not imply facts capable of being proved true or false." *Id.* § 566 cmt. b.

97. 750 F.2d 970, 979 (D.C. Cir. 1984), *cert. denied*, 471 U.S. 1127 (1985).

98. *Id.* at 979-85.

99. 497 U.S. 1, 19 (1990).

100. *Id.* at 3-4.

Maple Heights Public Schools, testified.[101] Following the hearing, the OHSAA censured Milkovich, placed his team on probation, and declared the team ineligible for the 1975 state tournament.[102] The parents of several members of the wrestling team promptly sued the OHSAA, claiming that they were denied due process in the OHSAA proceedings.[103] After a second hearing, in which Milkovich and Scott both again testified, the court overturned the OHSAA's orders.[104]

The next day, J. Theodore Diadiun, a sports columnist, wrote an article criticizing Milkovich's role in the altercation, as well as his testimony in the court proceeding.[105] The heading for his column stated, "Maple beat the law with the 'big lie.'"[106] The column included a passage stating that the message for Maple Heights students was "[i]f you get in a jam, lie your way out."[107] The column continued, asserting that "[a]nyone who attended the meet . . . [knew] in his heart that Milkovich and Scott lied at the hearing after each having given his solemn oath to tell the truth."[108] The tenor and language of the article clearly implied that Milkovich and Scott had perjured themselves, an indictable offense in the State of Ohio.[109]

Milkovich and Scott both sued the journalist, as well as his newspaper, for defamation, claiming that the published article accused them of perjury.[110] The Supreme Court of Ohio, applying *Ollman*'s four-factor analysis, held that Diadiun's column was constitutionally protected opinion.[111] The Court of Appeals of Ohio, in a separate action by Milkovich, concluded that it was bound by precedent and upheld a grant of summary judgment against Milkovich.[112] The Supreme Court of Ohio dismissed Milkovich's appeal for failing to raise a substantial constitutional issue, and Milkovich petitioned the U.S. Supreme Court.[113]

---

101. *Id.* at 4.
102. *Id.*
103. *Id.*
104. *Id.*
105. *Id.*
106. *Id.*
107. *Id.*
108. *Id.* at 5.
109. *Id.* at 21.
110. *Id.* at 6-7.
111. Scott v. News-Herald, 496 N.E.2d 699, 709 (Ohio 1986).
112. Milkovich v. News-Herald, 545 N.E.2d 1320, 1324 (Ohio Ct. App. 1989) (holding that, "as a matter of law, . . . the article in question was constitutionally protected opinion").
113. Milkovich v. Lorain Journal Co., 497 U.S. 1, 10 (1990).

1. The Supreme Court Opinion

The Supreme Court granted certiorari to consider constitutional issues raised by the Ohio courts.[114] The specific issues before the Court were (1) whether only statements of fact are actionable and (2) whether the distinction between opinion and fact should be determined under the *Ollman* four-factor test.[115]

After summarizing the constitutional evolution of defamation law, the Court referred to its famous dictum in *Gertz*,[116] which had been interpreted by numerous courts as providing First Amendment protection to any statement that could be labeled "opinion."[117] The Court rejected this view,[118] stating that such an interpretation would "ignore the fact that expressions of 'opinion' may often imply an assertion of objective fact."[119] Therefore, the opinion privilege should not immunize speakers from liability by prepending the magic words "in my opinion" to a statement.[120] The Court illustrated with the comment, "In my opinion Mayor Jones is a liar," which, although stated as opinion, could nevertheless be just as damaging to Jones' reputation as the assertion "Jones is a liar."[121] Such a statement is actionable because it implies unstated defamatory facts underlying the author's statement.[122]

The Court concluded that "the '"breathing space"' which '"freedoms of expression require in order to survive,"' is adequately secured by existing constitutional doctrine without the creation of an artificial dichotomy between 'opinion' and fact."[123] One such existing

---

114. *Id.*

115. *See id.* at 9.

116. Gertz v. Robert Welch, Inc., 418 U.S. 323, 339-40 (1974) ("Under the First Amendment there is no such thing as a false idea. However pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries but on the competition of other ideas. But there is no constitutional value in false statements of fact. Neither the intentional lie nor the careless error materially advances society's interest in 'uninhibited, robust, and wide-open' debate on public issues." (footnote omitted) (citation omitted)).

117. *See Milkovich*, 497 U.S. at 18 ("[T]his passage 'has become the opening salvo in all arguments for protection from defamation actions on the ground of opinion, even though the case did not remotely concern the question.'" (quoting Cianci v. New Times Publ'g Co., 639 F.2d 54, 61 (2d Cir. 1980))).

118. *Id.* ("[W]e do not think this passage from *Gertz* was intended to create a wholesale defamation exemption for anything that might be labeled 'opinion.'").

119. *Id.*

120. *See id.* at 20.

121. *Id.*

122. *Id.*

123. *Id.* at 19 (quoting Phila. Newspapers, Inc. v. Hepps, 475 U.S. 767, 772 (1986) (quoting N.Y. Times Co. v. Sullivan, 376 U.S. 254, 272 (1964))).

constitutional doctrine is the requirement that a plaintiff prove the falsity of a defamatory statement on a matter of public concern, as articulated in *Philadelphia Newspapers, Inc. v. Hepps*.[124] The Court reasoned that *Hepps* stands for the principle that "a statement of opinion relating to matters of public concern which does not contain a provably false factual connotation will receive full constitutional protection."[125]

A second protective constitutional doctrine identified by the Court was that of a line of Supreme Court cases that protects "loose, figurative, or hyperbolic" statements that cannot reasonably be understood as implying an assertion of objective fact about the plaintiff.[126] The special status of these types of expression derives from the constitutional protection provided for parody and other imaginative commentary by decisions such as *Hustler Magazine, Inc. v. Falwell*[127] and *Greenbelt Cooperative Publishing Ass'n, Inc. v. Bresler*,[128] rather than from any separate constitutional protection for opinion.[129]

The *Milkovich* Court rejected the dichotomy between fact and opinion, holding that the appropriate constitutional inquiry is not whether a statement constitutes fact or opinion, but whether it is capable of being proven true or false based on objective evidence.[130] The Court concluded that "a statement of opinion relating to matters of public concern which does not contain a provably false factual connotation will receive full constitutional protection."[131] To illustrate, the Court compared the statement, "In my opinion Mayor Jones is a liar," with the statement, "In my opinion Mayor Jones shows his abysmal ignorance by accepting the teachings of Marx and Lenin."[132] The former implies a verifiable fact and, thus, would be

124. 475 U.S. 767, 773 (1986).

125. *Milkovich*, 497 U.S. at 20.

126. *See id.* at 20-21; Hustler Magazine v. Falwell, 485 U.S. 46, 57 (1988) (stating that a parody of Rev. Falwell was not actionable because it "was not reasonably believable").

127. 485 U.S. 46 (1988).

128. 398 U.S. 6 (1970).

129. *Milkovich*, 497 U.S. at 20-21. The third rule identified by the Court was the fault requirements of *New York Times*, *Butts*, and *Gertz*. *Id.* at 20. The fourth rule was the appellate review standard established in *New York Times*, and reaffirmed in *Bose Corp. v. Consumers Union of U.S., Inc.*, 466 U.S. 485 (1984), which requires an appellate court to make an independent review of the finding of actual malice, when that standard is required. *Milkovich*, 497 U.S. at 21.

130. *Milkovich*, 497 U.S. at 20.

131. *Id.* (internal citation omitted).

132. *Id.*

actionable. The latter statement would receive full constitutional protection.

Thus, the *Milkovich* Court established verifiability as the sole criterion that determines the constitutional protection of a statement on a matter of public concern. Furthermore, a statement is verifiable in the constitutional sense only if its truth or falsity is based upon objectively determined facts.

Having set the analytical stage, the Court turned to the facts of the case before it with a two-step analysis. First, the Court ascertained the defamatory implication of the article about Milkovich and Scott published by the magazine.[133] The Court found that a reasonable fact finder could conclude that the article implied that petitioner Milkovich had perjured himself in a judicial proceeding.[134] Second, the Court determined the verifiability of the defamatory implication. The Court decided that it was indeed verifiable, reasoning:

> A determination of whether petitioner lied in this instance can be made on a core of objective evidence by comparing, *inter alia*, petitioner's testimony before the OHSAA board with his subsequent testimony before the trial court. . . . "Whether or not [petitioner] did indeed perjure himself is certainly verifiable . . . with evidence adduced from the transcripts and witnesses present at the hearing. Unlike a subjective assertion, the averred defamatory language is an articulation of an objectively verifiable event."[135]

The Court reversed and remanded the case, declaring that its decision struck an appropriate balance between the rights and guarantees of the First Amendment and the social values protected by the law of defamation.[136]

III. MALEVOLENT SOFTWARE

Malevolent software is a term for computer code that is designed to disrupt the operation of a computer system. The most common of these rogue programs are the computer virus and its common variant, the "worm."[137] Other forms of malicious software include so-called "logic bombs,"[138] "Trojan horses,"[139] and "trap doors."[140]

---

133. *Id.* at 21-22.

134. *Id.*

135. *Id.* (quoting Scott v. News-Herald, 496 N.E.2d 699, 707 (Ohio 1986)).

136. *Id.* at 22-23.

137. ED SKOUDIS & LENNY ZELTSER, MALWARE: FIGHTING MALICIOUS CODE 13-15 (2004).

138. A logic bomb is "[a] section of code, preprogrammed into a larger program, that waits for some trigger event to perform some damaging function. Logic bombs do not

A computer virus can be described as a series of instructions (a program) that: (1) infects a host program by attaching itself to the host; (2) executes when the host is executed; and (3) spreads by cloning itself, or part of itself, and attaching the clones to other host programs. In addition, many viruses have a so-called "payload," capable of harmful side effects, such as deleting, stealing, or modifying digital information.[141] As the definition suggests, a typical computer virus consists of three basic modules or mechanisms, namely an infection module, payload trigger, and payload.

A. Infection Module

The infection module enables a virus to reproduce and attach copies of itself onto target hosts. This mechanism is the most salient technical property of a computer virus.[142] The first task of the infection mechanism is to locate a prospective host program. Once a suitable host is found, the virus may take precautions, such as checking whether the host has already been infected.[143] The virus then installs a copy of itself on the host.[144] Once settled, the virus may

---

reproduce and so are not viral, but a virus may contain a logic bomb as a payload." DAVID HARLEY ET AL., VIRUSES REVEALED: UNDERSTAND AND COUNTER MALICIOUS SOFTWARE 654 (2001); *see also* PETER SZOR, THE ART OF COMPUTER VIRUS RESEARCH AND DEFENSE 30 (2005).

139. A Trojan horse is a program that appears to be beneficial, but contains a harmful payload. SZOR, *supra* note 138, at 31.

140. A trapdoor, or backdoor, is a function built into a program or system to allow unauthorized access to the system. *Id.* at 32; *see* DENNING & DENNING, *supra* note 3, at 75-78.

141. SKOUDIS & ZELTSER, *supra* note 137, at 27. In his PhD dissertation, Dr. Cohen defined a virus simply as any program capable of self-reproduction. *See* COHEN, *supra* note 3, at 1-2. This definition appears overly general. A literal interpretation of the definition would classify even programs such as compilers and editors as viral. *See* DENNING & DENNING, *supra* note 3, at 75.

142. ROGUE PROGRAMS: VIRUSES, WORMS, AND TROJAN HORSES 247 (Lance J. Hoffman ed., 1990) ("The ability to propagate is essential to a virus program."); DENNING & DENNING, *supra* note 3, at 73-75; HARLEY ET AL., *supra* note 138, at 87 ([T]he *infection mechanism* . . . is the code that allows the virus to reproduce [and infect a target host], and *thus to be a virus.* (second emphasis added)).

143. Viruses known as sparse infectors may try to slow down the rate of infection to avoid detection, while fast infectors, on the other hand, may attempt to infect as many hosts as possible. *See* HARLEY ET AL., *supra* note 138, at 87.

144. There are three mechanisms through which a virus can infect a host program. A virus may attach itself to its host as a shell, as an add-on, or as intrusive code. A shell virus forms a layer ("shell") around the host code, so that the latter effectively becomes an internal subroutine of the virus. The host program is then replaced by a functionally equivalent program that includes the virus. The virus executes first, and then allows the host code to execute. Boot program viruses are typically shell viruses. Most viruses are of the add-on variety. They become part of the host by appending, or prepending, their code to

take steps to protect itself from detection by changing its form.[145] When the host program runs,[146] control is passed to the resident virus code, allowing it to execute. The executing virus repeats the infection cycle by automatically replicating itself and copying the newly created clones to other executable files on the system or network, and even across networks.[147]

A virus may infect a computer or a network through several possible points of entry, including an infected file downloaded from the Internet, a web browser, removable media such as writable compact discs and DVDs, infected files in shared directories, an infected e-mail attachment, or infected commercial shrinkwrapped software. Fast-spreading worms, such as "CodeRed" and "Blaster," infect new hosts by exploiting network security vulnerabilities.[148] Early viruses targeted the boot sectors of floppy disks, a trend that continued into the 1990s.[149] Now, viruses are increasingly transmitted through e-mail attachments.[150]

E-mail is the most widely used medium of exchanging files and sharing information, but has also become a convenient and efficient vehicle for virus and worm propagation. For instance, fast-spreading viruses, such as "ExploreZip" and "Melissa," exploited automatic mailing programs to spread within and across networks.[151]

---

the host code, without altering the host code. The viral code may alter the order of execution, allowing itself to execute first and then the host code. Macro viruses are typically add-on viruses. Intrusive viruses, in contrast, overwrite some or all of the host code, replacing it with its own code. *See* DENNING & DENNING, *supra* note 3, at 81; PHILIP FITES ET AL., THE COMPUTER VIRUS CRISIS 73-75 (2d ed. 1992).

145. The capability to change its form is know as polymorphism. To detect polymorphic viruses requires a more complex algorithm than simple pattern matching. *See* DENNING & DENNING, *supra* note 3, at 89; *see also* HARLEY ET AL., *supra* note 138, at 87-88.

146. The execution of a host may be triggered by human intervention, such as when a user double-clicks on an infected e-mail attachment. *See* SKOUDIS & ZELTSER, *supra* note 137, at 26-27.

147. *Id.* at 31-37.

148. *See* SZOR, *supra* note 138, at 365-421.

149. In 1996, for instance, approximately 9 percent of respondents to a national survey listed e-mail attachments as the means of infection of their most recent virus incident, while 71 percent put the blame on infected diskettes. *See* Larry Bridwell, *ICSA Labs 10th Annual Computer Virus Prevalence Survey*, at 15 tbl. 5 (2004), *available at* http://www.icsalabs.com/icsa/docs/html/library/whitepapers/VPS2004.pdf.

150. In 2004, the corresponding numbers were 92 percent for e-mail attachments and 0 percent for diskettes. *Id.*

151. *See* Andy Bisset & Geraldine Shipton, *Some Human Dimensions of Computer Virus Creation and Infection*, 52 INT'L J. HUM. COMPUTER STUD. 899, 902 (2000) (citing Richard Ford, *No Surprises in Melissa Land*, 18 COMPUTERS & SECURITY 300, 302 (1999)).

B. Payload

In addition to replicating and spreading, viruses are often programmed to perform specific harmful actions. The module that implements this functionality is known as the payload.[152] A payload can be programmed to perform destructive operations, such as corrupting, deleting, and stealing information.[153] A payload may also create a backdoor that allows unauthorized access to the infected machine.[154] Some payload effects are immediately obvious, such as a system crash, while others are subtle, such as transposition of numbers and alteration of decimal places.[155] Subtle effects tend to be dangerous because their presence may not be detected until substantial harm has been done. Payloads are often relatively harmless and do no more than entertain the user with a humorous message, musical tune, or graphical display.[156]

A payload is triggered when a specific condition is satisfied. Triggering conditions come in a variety of forms, such as a specified number of infections, a certain date, or specific time. For instance, the "Friday-the-13th" virus only activated its payload on dates with the cursed designation.[157] In the simplest case, a payload executes whenever the virus executes, without waiting for a trigger event. Viruses do not always have a payload module, but even viruses

---

152. JAN HRUSKA, COMPUTER VIRUSES AND ANTI-VIRUS WARFARE 17-18 (1990) (noting that in addition to self-replicating code, viruses often also contain a payload, which is capable of producing malicious side-effects); *see also* COHEN, *supra* note 3, at 8-15 (providing examples of malignant viruses and what they do); JOHN MCAFEE & COLIN HAYNES, COMPUTER VIRUSES, WORMS, DATA DIDDLERS, KILLER PROGRAMS AND OTHER THREATS TO YOUR SYSTEM: WHAT THEY ARE, HOW THEY WORK, AND HOW TO DEFEND YOUR PC, MAC, OR MAINFRAME 60-61 (1989) (discussing the types of disruptions viruses may cause).

153. HARLEY ET AL., *supra* note 138, at 97-98.

154. HARLEY ET AL., *supra* note 138, at 88-89; SKOUDIS & ZELTSER, *supra* note 137, at 27; Meiring de Villiers, *Computer Viruses and Civil Liability: A Conceptual Framework*, 40 TORT TRIAL & INS. PRAC. L.J. 123, 172 (2004) (discussing damage due to virus infection).

155. *See* MCAFEE & HAYNES, *supra* note 152, at 61; *see also* SZOR, *supra* note 138, at 302 (describing "data diddlers" as "viruses that do not destroy data all of a sudden in a very evident form . . . [but] slowly manipulate the data, such as the content of the hard disk").

156. *See, e.g.*, Eric J. Sinrod & William P. Reilly, *Cyber-Crimes: A Practical Approach to the Application of Federal Computer Crimes Laws*, 16 SANTA CLARA COMPUTER & HIGH TECH. L.J. 177, 218 (2000) (describing the W95.LoveSong.998 virus, designed to trigger a love song on a particular date).

157. *See id.* at 217 n.176. More recently, the first CodeRed worm alternated between continuing its infection cycle, remaining dormant, and attacking the official White House Web page, depending on the day of the month. *See* Meiring de Villiers, *Free Radicals in Cyberspace: Complex Liability Issues in Information Warfare*, 4 NW. J. TECH. & INTELL. PROP. 13, 13 (2005).

without a payload may harm their environment by consuming valuable computing resources.[158]

A "worm" is a special type of virus. It is similar to a virus in most respects, except that it does not need to attach itself to a host program to replicate and spread. Like viruses, worms often carry destructive payloads, but even without a destructive payload a fast-spreading worm can do significant harm by slowing down a system through the network traffic it generates.[159]

IV. THE ANATOMY OF A VIRUS ALERT

Courts resolve the truth or falsity of a defamatory statement by considering factors such as the context of the statement and the information on which it was based.[160] In the case of a virus alert, context and information depend on the technology involved. A virus alert is generated by a computer program and is based on an assessment of digital patterns in other programs. The truth or falsity of an alert depends on the properties of these technologies that generated the alert. This Part discusses the most commonly used anti-virus technologies and the mechanisms by which they generate virus alerts.

Technical anti-virus defenses come in four varieties: (1) signature scanners, (2) activity monitors, (3) integrity checkers, and

---

158. Viruses can cause economic losses by replicating and spreading, such as filling up available memory space, slowing down the execution of important programs, and locking keyboards. The Melissa virus, for instance, mailed copies of itself to everyone in the victim's e-mail address book, resulting in clogged e-mail servers and even system crashes. *See, e.g.*, FITES ET AL., *supra* note 144, at 23 (noting that the Christmas card virus stopped a major international mail system just by filling up all available storage capacity); HARLEY ET AL., *supra* note 138, at 88 ("[A] virus does not necessarily need to have either a trigger or a payload. A virus with a trigger and payload but no replication mechanism is not, in fact, a virus, but may well be described as a Trojan.").

159. *See generally* Schoch & Hupp, *supra* note 3, at 172 (discussing the benefits and deleterious effects worms may have on a computer system or network).

160. *See* Milkovich v. Lorain Journal Co., 497 U.S. 1, 21-22 (1990) (stating that it can be objectively verified whether an individual had perjured himself, by looking at evidence of contradictions in his testimony, trial transcripts, and the testimony of other witnesses); Boule v. Hutton, 138 F. Supp. 2d 491, 504 (S.D.N.Y. 2001) ("'[C]hallenged statements are not to be read in isolation, but must be perused as the average reader would against the whole apparent scope and intent of the writing.'" (quoting Celle v. Filipino Reporter Enters., Inc., 209 F.3d 163, 177 (2d Cir. 2000))); *see also Celle*, 209 F.3d at 177 ("'[T]he words are to be construed not with the close precision expected from lawyers and judges but as they would be read and understood *by the public to which they are addressed*.'" (quoting November v. Time Inc., 194 N.E.2d 126, 128 (N.Y. 1963))); Armstrong v. Simon & Schuster, Inc., 649 N.E.2d 825, 829 (N.Y. 1995) (stating that courts "must give the disputed language a fair reading in the context of the publication as a whole" (citation omitted)).

(4) heuristic techniques.[161] Scanners detect known viruses by identifying patterns that are unique to each virus strain.[162] Activity monitors look out for virus-like activity in a computer.[163] Integrity checkers sound an alarm when detecting suspicious modifications to computer files.[164] Heuristic techniques combine virus-specific scanning[165] with generic detection,[166] providing a significantly broadened range of virus detection.

A. Scanners

Scanners are the most widely used anti-virus defense.[167] A scanner reads executable programs and searches for the presence of virus patterns, known as "signatures."[168] A virus signature consists of patterns of hexadecimal digits embedded in the viral code that are unique to a particular virus strain.[169] These signatures are created by human experts, at institutions such as IBM's High Integrity Computing Laboratory, who scrutinize viral code and extract sections of code with unusual patterns.[170] The selected byte patterns are collected in a signature database and used in anti-virus scanners.[171] A scanner detects a virus in a program by comparing the program to its database of signatures and announcing a match as a possible virus.[172]

An ideal virus signature would give neither false negatives nor false positives. In other words, it would always identify the virus

---

161. *See* DENNING & DENNING, *supra* note 3, at 90-93; KEN DUNHAM, BIGELOW'S VIRUS TROUBLESHOOTING POCKET REFERENCE 78-83, 102-08 (2000); HARLEY ET AL., *supra* note 138, at 139-70; SKOUDIS & ZELTSER, *supra* note 137, at 51-64; SZOR, *supra* note 138, at 425-93.

162. *See* SZOR, *supra* note 138, at 426.

163. HARLEY ET AL., *supra* note 138, at 151.

164. *Id.* at 155.

165. Virus-specific technology, such as signature scanners, detect known viruses by indentifying patterns that are unique to each virus strain; it identifies the specific strain it has detected. HARLEY ET AL., *supra* note 138, at 151-52.

166. Generic anti-virus technology detects the presence of a virus by recognizing generic virus-like behavior, usually without identifying the particular strain. *Id.* Integrity checkers and activity monitors are generic detectors. *Id.*

167. *Id.* at 158 ("Scanners, particularly signature scanners, are currently the most popular of antiviral software.").

168. *See* SKOUDIS & ZELTSER, *supra* note 137, at 53-55.

169. HRUSKA, *supra* note 152, at 42.

170. *See* SKOUDIS & ZELTSER, *supra* note 137, at 53-54 ("The antivirus vendors collect virus specimens and 'fingerprint' them.").

171. Jeffrey O. Kephart & William C. Arnold, *Automatic Extraction of Computer Virus Signatures*, *in* PROCEEDINGS OF THE 4TH VIRUS BULLETIN INTERNATIONAL CONFERENCE 178 (R. Ford ed., 1994), *available at* http://www.research.ibm.com/antivirus/SciPapers/Kephart/VB94/vb94.html.

172. *See* SKOUDIS & ZELTSER, *supra* note 137, at 53-55.

when it is present and never trigger an alarm when it is not.[173] Although this ideal is unachievable in practice, anti-virus researchers pursue optimal solutions within practical constraints. For instance, the IBM High Integrity Computing Laboratory has developed an optimal statistical signature extraction technique that examines all sections of code in a virus and selects the byte strings that optimize the tradeoff between false positives and false negatives.[174]

Scanners are easy to use, but they are limited to detecting known signatures.[175] A scanner's signature database has to be updated continually as new viruses are discovered and their signatures catalogued, a burdensome requirement in an environment where new viruses appear daily. Modern anti-virus vendors have attempted to lighten the burden on users by distributing signature updates directly to their customers via the Internet.[176]

False negatives are rare when scanning for viruses with known signatures, but false positives may arise when a signature has been chosen imprudently.[177] For instance, a scan string selected as a signature for a given virus strain may also be present in benign objects.[178] This pattern may then match code that is actually a harmless component of a legitimate program. Furthermore, a short and simple pattern can be found too often in innocent software and produce many false positives. Viruses with longer and more complex patterns, on the other hand, will give fewer false positives, but at the expense of more false negatives.[179] As the number of known viruses grows, the scanning process will inevitably slow down as a larger set of possibilities has to be evaluated.[180]

---

173. HRUSKA, *supra* note 152, at 42. For short descriptions and hexadecimal patterns of selected known viruses, see *id.* at 43-52. *See also* Kephart et al., *supra* note 5, at 11 ("[T]he signature extractor must select a virus signature carefully to avoid both false negatives and false positives. That is, the signature must be found in every instance of the virus, and must almost never occur in uninfected programs.").

174. Kephart & Arnold, *supra* note 171, at 178-84.

175. SKOUDIS & ZELTSER, *supra* note 137, at 53-54.

176. *Id.* at 54-55.

177. HARLEY ET AL., *supra* note 138, at 77-78.

178. *Id.* at 576; ROBERT SLADE, ROBERT SLADE'S GUIDE TO COMPUTER VIRUSES: HOW TO AVOID THEM, HOW TO GET RID OF THEM, AND HOW TO GET HELP 215 (2d ed. 1996) (noting that false positives are comparatively rare in virus scanners, but can occur if the digital signature for a given virus is not well chosen).

179. DUNHAM, *supra* note 161, at 78-83; Jeffrey O. Kephart et al., *Fighting Computer Viruses*, SCI. AM. (Nov. 1997), *available at* http://vx.netlux.org/lib/ajk01.html; *see also* Sandeep Kumar & Eugene H. Spafford, *A Generic Virus Scanner in C++*, *in* PROCEEDINGS OF THE 8TH COMPUTER SECURITY APPLICATIONS CONFERENCE 6-8 (1992), *available at* http://vx.netlux.org/lib/aes04.html.

180. *See, e.g.*, Pete Lindstrom, THE HIDDEN COSTS OF VIRUS PROTECTION 5 (June 2003) *available at* http://cnscenter.future.co.kr/resource/rsc-center/vendor-wp/sopho/

B. Activity Monitors

Activity monitors are resident programs that monitor activities in a computer for behavior commonly associated with viruses.[181] Suspicious activities include operations such as attempts by a program to delete information and mass mail copies of itself. When suspicious activity is detected, the monitor may simply halt execution and alert the user, or take definite action to neutralize the activity.[182] Activity monitors, unlike scanners, do not need to know the signature of a virus to detect it. Their function is to recognize general suspicious behavior, not the precise identity of the culprit.

The greatest strength of activity monitors is their ability to detect unknown virus strains, but they also have significant weaknesses. They can only detect viruses that are actually executing, possibly after substantial harm has been done. Furthermore, a virus may execute before the monitor code does, and do harm before the monitor is able to activate and detect it.[183] Additionally, a virus may be programmed to alter monitor code on machines that do not have protection against such modification.[184]

A further weakness of activity monitors is the lack of unambiguous rules defining "suspicious" activity.[185] This may result in false alarms when an activity monitor picks up legitimate activities that resemble virus-like behavior.[186] Recurrent false alarms may ultimately lead users to ignore warnings from the monitor. False negatives may result when an activity monitor fails to recognize viral activity that does not fit the monitor's programmed definitions.[187]

C. Integrity Verification

An integrity verifier applies the electronic equivalent of a tamper-proof seal to protected programs and issues an alert when the

---

spire.pdf ("In this day of 80,000+ known viruses and frequent discovery of new ones, the size of the signature file can be large, particularly if the updates are sent out as cumulative ones. Large updates can clog the network pipelines . . . and reduce the frequency that an administrator will push them out to the end users.").

181. *See* HARLEY ET AL., *supra* note 138, at 151.

182. Kumar & Spafford, *supra* note 179, at 3-4.

183. *See* HARLEY ET AL., *supra* note 138, at 153-58.

184. *See id.*

185. *See id.*

186. Using one program to delete another, formatting a floppy disk, and boot sector changes resulting from upgrading the operating system are all "legitimate" operations that may trigger false alarms. *See* SLADE, *supra* note 178, at 40-41.

187. *See* HRUSKA, *supra* note 152, at 75.

seal has been broken, presumably by the intrusion of a virus. An integrity verification program generates a code, known as a "checksum," for protected files.[188] A checksum may, for instance, be an arithmetic calculation based on the total number of bytes in a file, the numerical value of the file size, and its creation date.[189] A checksum is periodically recomputed and compared to the original.[190] When a virus infects a file, it usually modifies the contents, resulting in a change in the checksum.[191] When the recomputed value does not match the original, the file is presumed to have been modified since the previous inspection and a warning is issued.[192]

The advantage of integrity checking is that it detects most instances of viral infection, as infection usually alters the target file. Its main drawback is that it tends to generate false alarms, as a file can change for "legitimate" reasons unrelated to virus infection.[193] Therefore, integrity checking software presents a high likelihood of a false positive, given the general difficulty of determining whether a program change is legitimate or due to a virus.[194] Integrity checking works best on static files, such as system utilities, but is an inappropriate technique for files that change frequently, such as Word documents.

Intelligent analysis of file changes may reduce the incidence of false positives. A sophisticated integrity checker may, for instance, take into account the nature and location of a file change in determining whether it is viral.[195] Most integrity checkers include the option to exclude certain files or directories from monitoring.[196]

---

188. *See* HARLEY ET AL., *supra* note 138, at 155-58.

189. ROBERT SLADE, DICTIONARY OF INFORMATION SECURITY 34 (2006).

190. *See* HARLEY ET AL., *supra* note 138, at 155-58.

191. DUNHAM, *supra* note 161, at 79; FITES ET AL., *supra* note 144, at 69-76, figs. 5.2-.5; *see* Kumar & Spafford, *supra* note 179, at 5-6.

192. *See* SKOUDIS & ZELTSER, *supra* note 137, at 58. Integrity verification procedures can be used in antivirus software to detect viral infection. *Id.* If a file has been inexplicably modified, then the file may be infected, and the antivirus program should take a closer look at it. *Id.*

193. *See* FITES ET AL., *supra* note 144, at 125; SKOUDIS & ZELTSER, *supra* note 137, at 58.

194. SLADE, *supra* note 178, at 157.

195. A Microsoft Word document can, for instance, be expected to change when a user edits it, but modification in the case of a macro is much more suspicious.

196. SKOUDIS & ZELTSER, *supra* note 137, at 54 ("The user can specify that all files should be scanned for [malevolent] code . . . . As a more efficient but less thorough alternative, the user can require that only file types most likely to harbor viruses . . . be scanned.").

D. Heuristic Detection

A fourth category of virus detectors uses heuristic detection methods. Heuristic rules solve complex problems "fairly well" and "fairly quickly," but are less than perfect.[197] Virus detection is an example of a complex problem that is amenable to heuristic solution. It has been proven mathematically that it is impossible to write a virus detection program that is capable of consistent perfect detection.[198] Heuristic virus detection methods accept such limitations and attempt to achieve a heuristic solution, namely a detection rate that is below the (unachievable) perfect rate but that represents an optimal tradeoff between detection accuracy, speed, and computational expense.[199]

Heuristics detect novel viruses by examining the structure and logic of executable code for evidence of virus-like behavior. Based on this examination, the program assesses the likelihood that the scrutinized program constitutes a virus by tallying up a score. The heuristic scanner examines a file, assigns a weight to each virus-like feature it encounters, and calculates a score based on the weights. If a score exceeds a certain threshold, the scanner classifies the program as malicious code and notifies the user. For instance, instructions to send an e-mail message with an attachment to every listing in an address book would add significantly to the score. Other high-scoring routines include capabilities to replicate, to hide from detection, and to execute some kind of payload.[200]

A heuristic assessment is necessarily less than perfect and will inevitably provide false positives and false negatives. A low scanner threshold will result in false alarms. A scanner with a threshold that is set too high, on the other hand, will fail to detect viruses that are malicious but that do not exactly match the unrealistically tight specifications, resulting in false negatives.[201] As in the case of activity

197. Francisco Fernandez, *Heuristic Engines*, *in* PROCEEDINGS OF THE 11TH INTL. VIRUS BULLETIN CONFERENCE 407-11 (2001).

198. Diomidis Spinellis, *Reliable Identification of Bounded-Length Viruses is NP-Complete*, 49 IEEE TRANSACTIONS ON INFO. THEORY 280, 282 (2003) (stating that theoretically perfect detection is in the general case undecidable, and for known viruses, NP-complete); Carey Nachenberg, *Future Imperfect*, VIRUS BULL., Aug. 1997, at 6; *see also* Fernandez, *supra* note 197, at 407-44; David M. Chess & Steve R. White, *An Undetectable Computer Virus*, IBM Thomas J. Watson Research Center, *available at* http://www.research.ibm.com/antivirus/SciPapers/VB2000DC.htm (last visited Nov. 5, 2007).

199. *See* HARLEY ET AL., *supra* note 138, at 159; SZOR, *supra* note 138, at 467.

200. *See* SZOR, *supra* note 138, at 472-74.

201. SKOUDIS & ZELTSER, *supra* note 137, at 56.

monitors, the term "suspicious" is ambiguous. Many legitimate programs, including even some anti-virus programs, perform operations that resemble virus-like behavior.[202] Nevertheless, state-of-the-art heuristic scanners achieve a 70 to 80 percent success rate in detecting unknown viruses.[203]

A heuristic scanner typically operates in two phases. The scanning algorithm first narrows the search by identifying the location most likely to contain a virus. It then analyzes the code from that location to determine its likely behavior upon execution. A static heuristic scanner compares the code from the "most likely" location to a database of byte sequences commonly associated with virus-like behavior. The algorithm then decides whether to classify the code as viral.[204]

A dynamic heuristic scanner uses central processing unit (CPU) emulation.[205] It loads suspect code into a virtual computer, emulates its execution, and monitors its behavior. Because it is only a virtual computer, virus-like behavior can be safely observed in what is essentially a laboratory setting, with no need to be concerned about real damage.[206] Although dynamic heuristics can be time-consuming due to the relatively slow CPU emulation process, they are sometimes superior to static heuristics. This will be the case when the suspect code is obscure and not easily recognizable as viral in its static state, but clearly reveals its viral nature in a dynamic state.

A major advantage of heuristic scanning is its ability to detect viruses before they execute and cause harm. Other generic anti-virus technologies, such as behavior monitoring and integrity checking, can only detect and eliminate a virus based on suspicious behavior, usually after execution.[207] Heuristic scanning is capable of detecting novel virus strains whose signatures have not yet been catalogued. Conventional scanners cannot detect such strains. Heuristic scanners are also capable of detecting polymorphic viruses, a complex virus

---

202. Fernandez, *supra* note 197, at 409 ("Many genuine programs use sequences of instructions that resemble those used by viruses. Programs that use low-level disk access methods, TSRs, encryption utilities, and even anti-virus packages can all, at times, carry out tasks that are performed by viruses.").

203. Nachenberg, *supra* note 198, at 7.

204. Kumar & Spafford, *supra* note 179, at 4-5 (discussion in "Detection by static analysis/policy adherence").

205. The CPU of a computer is responsible for data processing and computation. *See* 1 DAVID BENDER, COMPUTER LAW § 2.02, at 2-7 (1978 & Supp. 2007); *see also* HRUSKA, *supra* note 152, at 113 ("The CPU . . . is the 'heart' of every computer. It is the device which takes instructions from memory and executes them.").

206. Kumar & Spafford, *supra* note 179, at 4.

207. *See, e.g.*, SKOUDIS & ZELTSER, *supra* note 137, at 58 ("The main limitation of the integrity verification method is that it detects the infection only after it occurs.").

family that complicates detection by changing its signature from infection to infection.[208]

The explosive growth in new virus strains has made reliable detection and identification of individual strains very difficult and costly, making heuristics more important and increasingly prevalent.[209] Commercial heuristic scanners include IBM's AntiVirus boot scanner and Symantec's Bloodhound technology.[210]

V. TRUTH, FALSITY, AND VERIFIABILITY OF A VIRUS ALERT

The United States Supreme Court revolutionized its First Amendment defamation jurisprudence with decisions in *New York Times Co. v. Sullivan*,[211] *Gertz v. Robert Welch, Inc.*,[212] *Philadelphia Newspapers, Inc. v. Hepps*,[213] and *Milkovich v. Lorain Journal Co.*[214] These decisions raise two implications, namely that (1) the plaintiff must plead a defamatory statement of fact that is objectively verifiable as true or false; and (2) the plaintiff must prove the falsity of the defamatory statement with convincing clarity, while the defendant may prove the truthfulness of the statement as a defense. The main issues arising from these implications concern the truth, falsity, and verifiability of the alleged defamatory statement.

A. Truth and Falsity

A defamation plaintiff must plead and prove the falsity of the statement at issue. Absolute truth is a complete defense to a defamation charge,[215] but a defendant does not have to prove the

208. Polymorphic viruses have the ability to "mutate" by varying the code sequences written to target files. To detect such viruses requires a more complex algorithm than simple pattern matching. *See* DENNING & DENNING, *supra* note 3, at 89.

209. Nachenberg, *supra* note 198, at 9.

210. *See* Gerald Tesauro et al., *Neural Networks for Computer Virus Recognition*, IEEE Expert, Aug. 1996, at 1, 5-6; Symantec Security Update, Bloodhound.Packed, http://www.symantec.com/security_response/writeup.jsp?docid=2004-012015-2255-99 (last visited January 10, 2008).

211. 376 U.S. 254 (1964).

212. 418 U.S. 323 (1974).

213. 475 U.S. 767 (1986).

214. 497 U.S. 1, 19 (1990).

215. 1 SACK, *supra* note 24, § 3.3.2.1 ("The Supreme Court has not decided whether the Constitution permits liability for truthful speech that fails the 'public concern' test, is not contained in the media, or both. But open or not, the question is largely academic. Even if courts *may* impose such liability, in practice they do not."); 1 SMOLLA, *supra* note 21, § 5:10 ("The better view, however, is that the first amendment's protection of truth, like its protection of opinion, stands on its own footing, and is analytically distinct from fault rules. . . . Just as under the first amendment there is 'no such thing' as a false idea, there should

literal truth of the defamatory statement to prevail. An effective defense can rely on the "substantial truth" doctrine.[216] This doctrine states that "'[t]ruth' will protect the defendant from liability even if the precise literal truth of the defamatory statement cannot be established."[217] Minor inaccuracies are immaterial as long as the "gist" or "sting" of the statement is true, regardless of who has the burden of proof and what standard of proof applies.[218] Then-Judge Antonin Scalia, writing for the District of Columbia Circuit court, provided the following illustrative example in *Liberty Lobby, Inc. v. Anderson*: suppose a newspaper reports that a person has committed thirty-five burglaries, while he has actually committed only thirty-four.[219] Although the statement is factually incorrect, it would not be actionable. It is substantially true, because its gist can be justified, namely that the person is a habitual burglar.[220]

The Supreme Court later formulated the substantial truth test as whether the libel as published "would have [had] a different effect on the mind of the reader from that which the pleaded truth would have produced."[221] Falsehoods that do not harm the plaintiff's reputation more than the full and accurate truth are, therefore, not actionable. For instance, in Justice Scalia's illustration, it would

---

be 'no such thing' as liability for defamation for speaking the truth."); *see also* Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc., 472 U.S. 749, 783-84 (1985) (Brennan, J., dissenting) (arguing there is no constitutional basis for a distinction between press and non-press in defamation cases); First Nat'l Bank v. Bellotti, 435 U.S. 765, 801 (1978) (Burger, J., concurring) (discussing the press/non-press distinction). Lower courts routinely follow *Hepps* in non-media cases. *See* Burroughs v. FFP Operating Partners, L.P., 28 F.3d 543 (5th Cir. 1994).

216. Vachet v. Cent. Newspapers, Inc., 816 F.2d 313, 316 (7th Cir. 1987); Zerangue v. TSP Newspapers, Inc., 814 F.2d 1066, 1073 (5th Cir. 1987) ("Truth is a defense to libel. . . . A publication is also protected if it is 'substantially true,' *i.e.*, if it varies from the truth only in insignificant details or if its 'gist' or 'sting' is true." (citations omitted)); Guccione v. Hustler Magazine, Inc., 800 F.2d 298, 301 (2d Cir. 1986) ("'[S]ubstantial truth' suffices to defeat a charge of libel.").

217. 1 SMOLLA, *supra* note 21, § 5:14.

218. Masson v. New Yorker Magazine, Inc., 501 U.S. 496, 516-17 (1991) ("The common law of libel takes but one approach to the question of falsity, regardless of the form of the communication. It overlooks minor inaccuracies and concentrates upon substantial truth." (internal citations omitted)).

219. 746 F.2d 1563, 1568 n.6 (D.C. Cir. 1984).

220. Subsequent courts have adopted Judge Scalia's example in applying the substantial truth doctrine. *See* Moldea v. N.Y. Times Co., 22 F.3d 310, 319 (D.C. Cir. 1994).

221. *Masson*, 501 U.S. at 517 (citations omitted); *see also* Gomba v. McLaughlin, 504 P.2d 337, 339 (Colo. 1972) ("The question, a factual one, is whether there is a substantial difference between the allegedly libelous statement and the truth; or stated differently, whether the statement produces a different effect upon the reader than that which would be produced by the literal truth of the matter."). A statement is not substantially true if the sting of the statement is worse than the exact truth. *See* 1 SMOLLA, *supra* note 21, § 5:24.

make no difference to a listener whether a habitual burglar had committed thirty-four or thirty-five burglaries.

A plaintiff has considerable control over the focus of a court's substantial truth analysis. Under common law pleading rules, a plaintiff must allege the defamatory meaning of the defendant's statement, namely the aspect of the statement that harmed his reputation.[222] If the court determines that the statement is not capable of bearing the meaning asserted by the plaintiff, it will dismiss the complaint.[223] If, on the other hand, the court determines that the interpretation is reasonable, it will hand the issue to the jury.[224] The jury then determines whether the defamatory meaning of the statement was so understood by the recipient, either correctly or mistakenly, but reasonably.[225] If the defendant's language was vague, the jury must determine whether the claimed meaning was in fact the recipient's interpretation.[226] The jury must then decide the truth or falsity of the defamatory meaning.[227]

Golden Bear Distributing Systems of Texas, Inc. v. Chase Revel, Inc. illustrates adjudication of the substantial truth issue.[228] In the case, a magazine printed an article on the activities of separate companies, all of them operating in different states under the name "Golden Bear Distributing Systems."[229] The article reported a lawsuit for investment fraud brought against Golden Bear of California and

---

222. RESTATEMENT (SECOND) OF TORTS § 563 cmt. f (1977).

223. Franklin & Bussel, *supra* note 49, at 865.

224. KEETON ET AL., *supra* note 9, § 116, at 782-83 ("[I]t remains a question for the court whether the meaning claimed might reasonably be conveyed, and for the jury whether it was so understood." (footnotes omitted)).

225. RESTATEMENT (SECOND) OF TORTS § 563.

226. KEETON ET AL., *supra* note 9, § 116, at 781 ("If the language used is open to two meanings, as in the case of the French word 'cocotte,' which . . . signifies either a prostitute or a poached egg, it is for the jury to determine whether the defamatory sense was the one conveyed." (footnote omitted)).

227. Fields Found., Ltd. v. Christensen, 309 N.W.2d 125, 135 (Wis. Ct. App. 1981); RESTATEMENT (SECOND) OF TORTS § 617(b) ("Subject to the control of the court whenever the issue arises, the jury determines whether . . . (b) the matter was true or false . . . ."); *see also* Thomas Gibbons, *Defamation Reconsidered*, 16 OXFORD J. LEGAL STUD. 587, 606 (1996) ("The jury must decide by looking to the 'gist' or the 'sting' of the allegation, asking whether the discrepancy between the facts required to be justified and those proved to be true would make any difference to the judgment of reasonable people."); Lisa K. Snow, Note, *A Broader Approach to the Substantial Truth Defense*, 29 B.C. L. REV. 769, 785 (1988) ("The jury is the appropriate factfinder in this situation because the jury represents the average reader or listener. The jury is in the best position to determine whether the opprobrium attached to the alleged misstatement is similar to the opprobrium attached to the act actually committed.").

228. 708 F.2d 944 (5th Cir. 1983).

229. *Id.* at 946.

described legal difficulties plaguing Golden Bear of Utah.[230] The article also referred to the marketing strategy of Golden Bear of Texas, noting its similarity to strategies of the troubled Golden Bear franchises.[231] The article did not state that the Texas franchise was guilty, or even accused, of any wrongdoing.[232] However, when the article appeared, Golden Bear of Texas rapidly lost business and was forced into bankruptcy.[233] Golden Bear of Texas successfully sued the magazine for libel and was awarded damages at trial.[234] The judgment was affirmed by the Fifth Circuit on appeal.[235]

The Fifth Circuit observed that all of the individual statements in the magazine article concerning Golden Bear of California's legal problems and the reference to the marketing strategy of Golden Bear of Texas were literally true.[236] However, the court accepted the plaintiff's pleading that the magazine article falsely implied that Golden Bear of Texas had engaged in misconduct.[237] While the defendant argued that its article was substantially true, the court did not agree.[238] The factual allegations of the defendants, while literally true, did not justify the defamatory implication, namely that Golden Bear of Texas had engaged in misconduct similar to that of Golden Bear of California.[239] The plaintiff succeeded in proving the falsity of the defamatory gist of defendant's communication, while the defendant's pleaded truth failed to justify it.[240]

B. Substantial Truth Analysis of a Virus Alert

The aim of this subsection is to define and analyze the legal meaning of the concept "substantial truth of a virus alert." A virus alert literally states that "a program is infected with a specific type of malevolent code."[241] Its truth as a technical statement is not

---

230. *Id.* at 947.
231. *Id.*
232. *Id.* at 948.
233. *Id.* at 947.
234. *Id.* at 946.
235. *Id.* at 952.
236. *Id.* at 948. See also 1 SMOLLA, *supra* note 21, § 5:20, for a discussion of *Golden Bear*.
237. *Golden Bear*, 708 F.2d at 949.
238. *Id.*
239. *Id.*
240. *Id.*
241. SLADE, *supra* note 178, at 8 (Defining an alert as "notification that an *event* or *incident* has occurred."). The event or incident announced by anti-virus software is that a virus has been found in the files examined. *See* JOSE NAZARIO, DEFENSE AND DETECTION STRATEGIES AGAINST INTERNET WORMS 175 (2003) (A signature analysis approach to worm

controversial. Its truth as a defamatory statement and in a constitutional sense, however, is a legal concept that requires analysis of the defamatory meaning of the alert. The analysis suggests that a virus alert is substantially true if, and only if, the detected object is capable of executing an infection module. This conclusion is based on the following reasoning:

> 1. A virus alert charges the plaintiff with compromising the information security of the computing environment where its software is implemented. Therefore, the reputational interest at stake in a virus alert is the plaintiff's reputation for the security of its software product.
>
> 2. The defamatory meaning of a virus alert is that aspect of the alert that threatens the plaintiff's reputation. The reproductive capability of a virus is the essence of its threat to information security and, indirectly, the plaintiff's reputation. Therefore, proof of a reproductive capability is sufficient to prove the truth of the defamatory meaning—the substantial truthfulness—of a virus alert.
>
> 3. A statement is substantially true if it is factually similar to the proven truth and differs from the truth by no more than insubstantial details. The Supreme Court's test of substantiality is whether an allegation "would have [had] a different effect on the mind of the reader from that which the pleaded truth would have produced."[242] The presence of an executable infection module in malevolent software is material to the average computer user because such a module creates a unique and serious type of risk to information security, a risk that malevolent code without such a module does not possess. Therefore, proof of the truth of a virus alert must include proof of an executable infection module. In other words, a virus alert is substantially true only if the detected object has a reproductive capability.
>
> 4. In conclusion, the presence of an executable infection module in the detected object is a necessary and sufficient condition for a virus alert to be substantially true.

1. Plaintiff's Reputational Interest

A software vendor's most vital intangible asset is its reputation for secure software, especially in the security-conscious environment following the terrorist attacks of September 11, 2001. A vendor of software, especially software destined for the networks of the national

---

detection compares a list of malevolent patterns against network traffic and issues an alert when a match is found.). Some anti-virus products identify a detected virus by name, while others make a generic identification. *See* HARLEY ET AL., *supra* note 138, at 143 ("Virus-specific software takes the approach, 'I have identified virus X.'"); *id.* at 144 ("Generic detection software deduces the presence of a virus from environmental anomalies. It doesn't identify a specific virus by name."). A virus alert therefore states that a program is infected with a specific type of malevolent code, namely a virus or a worm. A virus alert may in some circumstances be even more specific and identify the virus or worm by name.

242. Masson v. New Yorker Magazine, Inc., 501 U.S. 496, 517 (1991) (citation omitted).

information infrastructure[243] and other sensitive applications, is dependent on a reputation for security to survive and remain in business.[244] The most significant threat to information security is the proliferation of computer viruses and worms on the Internet.[245] Thus, a vendor's reputation for secure software may be irreparably harmed by a virus alert that indicates the presence of such malevolent code in its software product.

Modern information security has three basic components: (1) confidentiality, (2) integrity, and (3) availability.[246] Confidentiality refers to the prevention of unauthorized access to sensitive information.[247] Integrity refers to the protection of digital data from unauthorized change, such as corruption or deletion.[248] Availability refers to procedures and safeguards ensuring that authorized users have access to information when needed and in a convenient format.[249] A computer virus threatens all components of information security through its capability to replicate and spread.[250] The infection module also serves to export and multiply the effect of a payload (if the virus has a payload). Most viruses do not have a payload, though, and a

---

243. The national information infrastructure is an interrelated system of computer and communication networks that control and coordinate essential infrastructures, such as water supplies, banking and financial services, telecommunications services, and electrical power. It also includes computer networks that coordinate and control military communications and logistics. The private sector plays a dominant role in the critical information infrastructure. Most infrastructures are owned by the private sector, and the Defense Information Systems Agency depends heavily on commercial communication networks. *See* GLOBAL ORGANIZED CRIME PROJECT, CTR. FOR STRATEGIC & INT'L STUDIES, CYBERCRIME . . . CYBERTERRORISM . . . CYBERWARFARE . . . : AVERTING AN ELECTRONIC WATERLOO, at xiv-xv (1998).

244. *See, e.g.*, RICK LEHTINEN ET AL., COMPUTER SECURITY BASICS 27 (2d ed. 2006) ("In the 2000s, particularly after the attacks of 9/11, security took on a serious tone. Corporations and government alike became more willing to make security an integral part of their products and their jobs."); *see also id.* ("The challenge of this decade will be to consolidate what we have learned—to build computer security into our products and our daily routines, to protect data without unnecessarily impeding our ability to access it, and to make sure that both security products and government and industry standards grow to meet the ever-increasing scope and challenges of technology.").

245. *See* SKOUDIS & ZELTSER, *supra* note 137, at 25-27.

246. *See* MATT BISHOP, INTRODUCTION TO COMPUTER SECURITY 1-6 (2005); RICK LEHTINEN ET AL., *supra* note 244, at 9.

247. *See* RICK LEHTINEN ET AL., *supra* note 244, at 9 ("Data is confidential if it stays obscure to all but those authorized to use it.").

248. *See id.* at 11.

249. *See id.* at 12.

250. *See* HARLEY ET AL., *supra* note 138, at 97 ("Direct damage can be considered in terms of the classic tripartite security model (Availability, Integrity, Confidentiality). Viruses . . . have an impact across all three areas described by this model, as well as other areas, such as accountability.").

payload is not essential to be classified as a virus[251] or to threaten information security.[252]

a. Viral Code Threatens the Confidentiality of Information

A virus or worm can be programmed to access and steal confidential information on a system.[253] The "W32/Bugbear@mm" (Bugbear) family of viruses, for instance, was designed to exploit vulnerabilities in the Outlook e-mail program to gain access to machines, steal confidential information using a keylogging function, and interfere with antivirus software.[254] It also created a backdoor for hackers to take over the machine and misappropriate passwords and confidential financial information.[255] Some members of the Bugbear family specifically targeted financial institutions.[256]

Viruses and worms often use spoofed e-mail and Web sites to deceive users into disclosing confidential information, a technique known as "phishing."[257] The "W32/Mimail.I@mm" worm, for instance, displayed dialogues, purportedly from the online-payment service PayPal, requesting financial information from unwitting users. The stolen information was then encrypted and transmitted to the attacker.[258]

---

251. *See, e.g.*, *id.* at 7 ("[O]nly the presence of the infection mechanism is mandatory if the program is to be defined as viral: payload and trigger are optional.").

252. This section argues that a virus is capable of threatening all aspects of information security through its infection module alone.

253. *See* SKOUDIS & ZELTSER, *supra* note 137, at 3-4 (noting that a virus could steal files from your machine, especially sensitive ones containing personal, financial, or other sensitive information). Viruses also monitor user keystrokes and transmit information about the user's computing habits, websites visited, and financial information to the attacker. *See id.* at 3.

254. *Virus Makes Unwelcome Return*, BBC NEWS, June 5, 2003, http://news.bbc.co.uk/1/hi/technology/2965924.stm.

255. *Id.*

256. *Id.*; *see also* Gregg Keizer, *Virus Posing as Microsoft E-Mail Spreads Fast*, INFORMATIONWEEK, Sept. 19, 2003, http://www.informationweek.com/story/showArticle.jhtml?articleID=15000134 (describing a fast-spreading worm which attempts to steal confidential information from infected systems); Chariot Security Information, http://www.chariot.net.au/viruslist.php?page=101031&v=1 (last visited Nov. 7, 2007) (describing the W32.Sobig and Klez worms, which have been programmed to steal confidential information on infected machines).

257. SLADE, *supra* note 189, at 142.

258. *See* SZOR, *supra* note 138, at 308-09.

b. Malicious Code Threatens the Integrity of Information

Viral payloads can be programmed to delete, modify, or corrupt information on infected computers.[259] In January 2003, a young Welshman, Simon Vallor, was sentenced to two years imprisonment for releasing fast-spreading viruses via e-mail that were designed to corrupt data on the hard drives of infected computers.[260] Viruses often corrupt information by replicating and spreading alone, without the help of a payload. According to Peter Szor, a leading anti-virus researcher, "[v]irus replication . . . has many side effects. This includes the possibility of accidental data loss when the machine crashes due to a bug in the virus code or accidental overwriting of a part of the disk with relevant data. Virus researchers call this kind of virus a *no payload* virus."[261]

c. Malicious Code Threatens the Availability of Information

Fast-spreading viruses make infected systems unavailable to legitimate users by monopolizing valuable computational resources.[262] For example, a recent denial of service attack on the Port of Houston made crucial navigating data on the port's Web service temporarily unavailable to shipping pilots and mooring companies, creating substantial collision and other risks.[263] The Internet worm "W32/CodeRed" and its successors were deployed to exploit a vulnerability in Microsoft's Internet Information Services (IIS) web servers[264] to create a global denial of service effect on the Internet.[265]

---

259. *See Computer Virus*, COLUMBIA ENCYCLOPEDIA (6th ed. 2005), *available at* http://www.bartleby.com/65/co/computer-vir.html ("Although some viruses are merely disruptive, others can destroy or corrupt data or cause an operating system or applications program to malfunction.").

260. *Computer Virus Author Jailed,* BBC NEWS, Jan. 21, 2003, http://news.bbc.co.uk/2/hi/uk_news/wales/2678773.stm.

261. *See* SZOR, *supra* note 138, at 296-97.

262. *See* HARLEY ET AL., *supra* note 138, at 94 ("Network and mail viral programs carry, in a sense, their own payloads. The reproduction of the programs themselves uses the resources of the hosts affected and, in the cases of both the Morris Internet and CHRISTMA worms, went so far as to deny service by using all available computing or communications resources."); GREG HOGLUND & GARY MCGRAW, EXPLOITING SOFTWARE: HOW TO BREAK CODE 20 (2004) ("Worms allow an attacker to 'carpet bomb' a network in an unbridled exploration that attempts to exploit a given vulnerability as widely as possible. This amplifies the overall effect of an attack and achieves results that could never be obtained by manually hacking one machine at a time."); *see also* SZOR, *supra* note 138, at 306-07 (discussing denial of service attacks).

263. *See* Steve Gibson, *The Strange Tale of the Denial of Service Attacks Against GRC.COM*, Sept. 17, 2005, http://grc.com/dos/grcdos.htm.

264. The attacks occurred shortly after Microsoft had discovered the vulnerability and issued a patch to fix it. CERT, *A Very Real and Present Threat to the Internet: July 31*

The "W32/Slammer" (Slammer) worm overloaded Internet routers and slowed down networks worldwide, making it difficult to use e-mail. The paralyzing effect of Slammer on the Internet also caused ATM failures and interfered with elections.[266] The "Sasser" worm scanned so aggressively for new target computers that it caused networks to become congested and slow down. In Australia, Sasser disrupted Railcorp trains and brought down the computer system of Westpac Bank, a major Australian financial institution.[267] In the UK, Sasser caused flight delays and brought down the computerized mapping systems of several coastguard stations.[268]

In conclusion, the reproductive capability of a virus is the essence of its threat to information security and, indirectly, the plaintiff's reputation. Therefore, proof of a reproductive capability is sufficient to prove the truth of the defamatory meaning—the substantial truthfulness—of a virus alert.

2. Evidentiary Precision

Proof of the truthfulness of a defamatory allegation must be as precise and specific as the allegation itself.[269] An allegation that a plaintiff embezzled money cannot be justified by proving that the plaintiff breached a fiduciary duty,[270] and a charge that a plaintiff committed a burglary cannot be justified by proving that the plaintiff committed a murder.[271]

The common law position seems unduly strict. A defendant may, with apparent justification, argue that an allegation that the plaintiff had embezzled money does not harm the plaintiff's reputation substantially more than the exact truth, namely that the plaintiff had

---

Deadline for Action, July 29, 2001, http://www.cert.org/congressional_testimony/CRannounce.html.

265. *See* Kevin J. Houle & George M. Weaver, *Trends in Denial of Service Attack Technology*, at 19 (2001), http://www.cert.org/archive/pdf/DoS_trends.pdf.

266. SZOR, *supra* note 138, at 306.

267. *Worm Brings Down Coastguard PCs*, BBC NEWS, May 4, 2004, http://news.bbc.co.uk/2/hi/technology/3682803.stm.

268. *Id.*

269. *See* 1 SMOLLA, *supra* note 21, § 5:19 ("When the defamatory allegation is narrow and specific, the evidence of truth must more strictly conform to the allegation." (footnote omitted)).

270. *See, e.g.*, Roper v. Mabry, 551 P.2d 1381, 1395 (Wash. Ct. App. 1976). "[T]he defense to this defamation action is proof of the truth of the statements that Mr. Roper is a 'thief' or 'embezzler', [sic] and not proof that he breached a fiduciary duty." *Id.* at 1385.

271. *See, e.g.*, Barlow v. Int'l Harvester Co., 522 P.2d 1102, 1112 (Idaho 1974) (finding that proof of one criminal act does not justify a charge alleging commission of a different crime).

breached a fiduciary duty. However, Professor Smolla explains that "[t]he relative strictness of the common law position on substantial truth when specific defamatory charges are made can be justified on the grounds that more detailed charges of misconduct often tend to create greater reputational injury, simply because the existence of detail tends to lend credibility to the accusation."[272] A defamer who uses specificity to strengthen the credibility of his story must pay the price, namely be required to prove the truth with evidence as precise as the allegation itself.[273]

Courts distinguish between inaccuracies where the allegation differs factually from the truth and inaccuracies where the allegation is factually similar to the truth, but errs in insubstantial details.[274] A statement would be substantially true if it is factually similar to the proven truth and differs from the truth by no more than insubstantial details.[275] An allegation that a plaintiff embezzled money is substantially false if, in fact, the plaintiff only breached a fiduciary duty. Justice Scalia provided an example of a newspaper report that a person committed thirty-five burglaries, while he actually committed only thirty-four.[276] The report is substantially true because it differs from the factual truth only in an insubstantial detail.

The Supreme Court has analyzed the substance of a communication by looking at the mental impact of the communication on the average recipient. The test is whether the allegation "would have [had] a different effect on the mind of the reader from that which the pleaded truth would have produced."[277] This subsection analyzes the doctrines governing substantial truth in the context of a virus alert.

Consider a virus alert based on a type of malevolent code without an executable infection module, such as a logic bomb.[278] At a

---

272. 1 SMOLLA, *supra* note 21, § 5:20.

273. The defamation defendant does not, of course, bear the burden of proof of truth, but may choose to plead a truth defense.

274. *See* 1 SMOLLA, *supra* note 21, § 5:20; *see also* 1 DAN B. DOBBS, THE LAW OF TORTS 1148 (2001 & Supp. 2007) ("[I]f (a) the publication states facts similar to the truth and (b) the sting of the publication is substantially equivalent to the sting of the truth, the truth defense should ordinarily apply.").

275. *See* 1 SMOLLA, *supra* note 21, § 5:20.

276. Liberty Lobby, Inc. v. Anderson, 746 F.2d 1563, 1568 n.6 (D.C. Cir. 1984).

277. *See* Masson v. New Yorker Magazine, Inc., 501 U.S. 496, 517 (1991) (citation omitted); Chung v. Better Health Plan, No. 96 CIV 7310, 1997 U.S. Dist. LEXIS 9627, at *5 (S.D.N.Y. July 8, 1997).

278. A logic bomb lies dormant until an event, such as a pre-programmed date or time is reached. It then activates and executes a payload, but it cannot replicate and spread. *See* SZOR, *supra* note 138, at 30. A logic bomb always has a payload, but unlike a virus, it has no infection module. *Id.* A logic bomb has been compared to a real-world

high level of abstraction, a virus alert based on a logic bomb makes a factually correct statement, namely that the detected object constitutes malevolent code. Therefore, the virus alert is factually similar to the proven truth, but the specificity of the alert communicates additional information. When an alert identifies an object as a virus, it implies that the object not only constitutes malevolent code, but that it also contains an executable infection module. If the presence of an executable infection module in malevolent code is material to the average computer user, the virus alert would be substantially false if it identifies an object without such a module as viral.[279]

The harm threatened by malevolent code without a reproductive capability, such as a logic bomb, is limited by its inability to spread beyond the system where it was planted. A virus, on the other hand, threatens the confidentiality, integrity, and availability of information far beyond its origin. Dr. Fred Cohen provides a dramatic illustration: "Sitting at my Unix-based computer in Hudson, Ohio, I could launch a virus and reasonably expect it to spread through 40% of the Unix-based computers in the world in a matter of days. That's dramatically different from what we were dealing with before viruses."[280] Dr. Cohen's statement was published more than a decade ago. Today, viruses spread much faster, and there is every indication that virus transmission will continue to accelerate. The 2003 ICSA report remarks, for instance, that, while it took the early file viruses months to years to spread widely, subsequent macro viruses took weeks to months, mass mailers took days, "Code Red" took approximately 12 hours, and "Klez" spread around the world in 2.5

---

landmine. *See* TECH-FAQ, *What is a Logic Bomb?*, http:www.tech-faq.com/logic-bomb.shtml (last visited Nov. 7, 2007).

279. *See* Currier v. W. Newspaper, Inc., 855 P.2d 1351, 1354 (Ariz. 1993) ("A technically false statement may nonetheless be considered substantially true if, viewed 'through the eyes of the average reader,' it differs from the truth 'only in insignificant details.'").

280. COHEN, *supra* note 3, at 25; *see* CLIVE GRINGRAS, THE LAWS OF THE INTERNET 58 (1997) ("A computer file harbouring [sic] a virus can, in a matter of hours, spread across continents, damaging data and programs without reprieve."); *see also* FITES ET AL., *supra* note 144, at 21-22 (discussing the history of computer virus programs); Bradley S. Davis, *It's Virus Season Again, Has Your Computer Been Vaccinated? A Survey of Computer Crime Legislation as a Response to Malevolent Software*, 72 WASH. U. L.Q. 411, 437 n. 225 ("[A] user whose computer was infected could connect to an international network such as the Internet and upload a file onto the network that contained a strain of malevolent software. If the software was not detected by a scanning system . . . on the host computer, infection could spread throughout the Internet through this simple exchange of data." (citation omitted)).

hours.[281] Whereas code such as a logic bomb can destroy data worth *D*, releasing a virus to do the same job can cause that same harm several times over by spreading into *N* number of systems, causing damage of magnitude *N* * *D*, where *N* can be very large. This distinction is clearly material to computer users.

Thus, a virus alert requires proof of a reproductive capability to be substantially true. Proof that an object is a logic bomb does not justify calling it a virus, just as proof that a plaintiff committed a single homicide does not justify calling him a mass murderer.[282] Identifying an object without a reproductive capability as a virus materially mischaracterizes it and significantly misstates the nature, as well as the degree, of harm of which it is capable. Furthermore, the specificity of a virus alert, as a warning that implies a risk that could escalate into an electronic tsunami as opposed to a localized threat, strengthens the credibility and impact of the communication. Under common law evidentiary standards and the Supreme Court's mental impact test,[283] proof of the truth of a virus alert should therefore include proof of an executable infection module.

The analysis in this subsection has provided two major conclusions. First, proof of a reproductive capability is sufficient to prove the truth of the defamatory meaning—the substantial truthfulness—of a virus alert. Second, a virus alert is true only if the detected object contains an executable infection module. Therefore, proof of a reproductive capability is necessary to prove the substantial truthfulness of a virus alert.

Thus, a virus alert is substantially true if, and only if, the detected object contains an executable infection module. Put differently, the presence of an executable infection module is necessary and sufficient for a virus alert to be substantially true. A plaintiff may prove the falsity of a virus alert by demonstrating that the detected object either does not have an infection module or that it has an infection module that cannot execute, perhaps due to a programming or logical error in the module's code.

---

281. ICSA Labs 9th Annual Computer Virus Prevalence Survey (2003), http://www.securitymanagement.com/archive/library/ICSA_Virus 0604.pdf.

282. *See* Barlow v. Int'l Harvester, Inc., 522 P.2d 1102, 1112 (Idaho 1974) (finding that proof that plaintiff had committed one crime does not justify a false allegation that he had committed a different crime); KEETON ET AL., *supra* note 9, at 841 ("[A defendant] may not avoid liability by proving that the imputation was true in part, or, if the charge is one of persistent misconduct, by showing that it was true in a single instance." (footnotes omitted)).

283. Masson v. New Yorker Magazine, Inc., 501 U.S. 496, 517 (1991).

C. Forensic Proof of Truth and Falsity

Courts resolve the truth or falsity of a defamatory statement by considering factors such as the context of the statement and the information on which it was based.[284] In the case of a virus alert, context and information are creatures of technology. A virus alert is generated by a computer program and is based on an assessment of digital patterns in other programs.[285] The truth or falsity of an alert therefore depends on the properties of the technologies that generated the alert.

A virus alert is substantially true only if the identified code is capable of executing an infection module. This capability can often be conclusively verified by analyzing the code. Analysis of the logic of the host program, the viral code, and its infection module, as well as the absence of programming and logical errors, may reveal that the host will be run, control will be passed to the viral code, and the infection module will be triggered. In such a technically complex but uncontroversial case, the virus alert would be provably true and the defendant should prevail on a truth defense.

A more complicated situation arises when the detected virus cannot execute on the system in which it was found, even though it could execute on another system. Such viruses are known as "latent" or "dormant."[286] "For example, a PC-specific program infected by a PC-specific file virus cannot normally be executed on a UNIX server or a Macintosh."[287] It may nevertheless be found in these "foreign" environments, perhaps in an FTP directory[288] or as part of an e-mail attachment. The dormant virus may later "wake up" when it is transferred to a system on which it could execute, perhaps by e-mail or through file sharing. This kind of transmission is known as "heterogeneous virus transmission."[289]

---

284. *See* Milkovich v. Lorain Journal Co., 497 U.S. 1, 21-22 (1990) (stating that it can be objectively verified whether an individual had perjured himself, by looking at evidence of contradictions in his testimony, trial transcripts, and the testimony of other witnesses); Boule v. Hutton, 138 F. Supp. 2d 491, 504 (S.D.N.Y. 2001) (stating that alleged defamatory statements should not be read in isolation, but should be reviewed within whole context of the publication as the average, reasonable, intended reader would (citing Celle v. Filipino Rptr. Enters. Inc., 209 F.3d 163, 177 (2d Cir. 2000))).

285. *See supra* Part IV.

286. HARLEY ET AL., *supra* note 138, at 8.

287. *Id.* at 9.

288. An FTP, or File Transfer Protocol, is a program that connects two computers so that files can be transferred between them.

289. HARLEY ET AL., *supra* note 138, at 144.

The truth or falsity of a virus alert based on a dormant virus may be controversial. If heterogeneous transmission of the dormant virus were possible,[290] its infection module would, in principle, be executable and the virus alert would be substantially true. The state—transfer to an environment where it can execute—can be interpreted as a triggering condition that has to be satisfied to execute the infection module. If this trigger can be satisfied, the alert would be substantially true. The dormant virus cannot do harm in its current environment, but it is nevertheless a security threat. By analogy, a firearm may justifiably be described as "dangerous," even if currently in possession of a responsible person, if it could easily fall into unsafe hands.[291]

A dormant virus should be distinguished from a virus that cannot execute because of a logical or programming defect in its code. Neither the dormant virus of the previous example, nor the defective "virus," can execute in their current states, but both can be transformed into a state where they can execute. The dormant virus can be transferred to a new environment and the defective virus can be debugged. If an alert based on the dormant virus is substantially true, the superficial similarity of these situations may seem to suggest that an alert based on the defective virus must also be substantially true. However, it ignores the fact that the truth of an alert must be evaluated with respect to the object on which the alert was based, and as of the time when the alert was communicated.[292] In the dormant case, the virus on which the alert was based is identical to the virus that can execute in the new environment. Therefore, the virus alert is correct when it identifies the dormant object as viral. In the case of the defective virus, in contrast, the virus on which the alert was based is not the same as the virus that could eventually execute. Furthermore, at the time the alert was communicated, the object of the communication was not executable and, thus, not a virus. The executable virus is a corrected version of the defective object on which

290. A PC-specific program infected by a PC-specific file virus resident in an e-mail attachment, on a Macintosh, for instance, may be transmitted to a PC where it could execute.

291. *See, e.g.*, BISHOP, *supra* note 246, at 4 ("A *threat* is a potential violation of security. The violation need not actually occur for there to be a threat. The fact that the violation *might* occur means that those actions that could cause it to occur must be guarded against (or prepared for.) Those actions are called *attacks*. Those who execute such actions, or cause them to be executed, are called *attackers*.").

292. *See* RESTATEMENT (SECOND) OF TORTS § 581A cmt. g (1977) ("The truth of a defamatory imputation of fact must be determined as of the time of the defamatory publication.").

the alert was based. A virus alert based on the original defective virus would therefore be false.

False positives are comparatively rare in virus scanners, but they may occur if a virus signature is not well chosen. For instance, a signature that also occurs in legitimate code may cause a scanner to misdiagnose a program as infected.[293] A type of false positive known as a "ghost positive" is generated when remnants of a virus are incorrectly detected and reported as viral.[294] Ghost positives may occur when an antivirus program attempts to remove a virus from an infected file but leaves part of the virus code intact. The remnant code may contain a virus signature, even though the disembodied remnant cannot execute. Another anti-virus program may subsequently scan the file, detect the remnant, and report it as viral. Ghost positives also occur when a computer user installs two or more scanners simultaneously on the same computer. One of the scanners may fail to encrypt its virus signatures and store them in plain text, exactly in the format in which they would appear in an infected file.[295] The other scanner may then identify these signatures as a viral presence in the computer.[296] The detected objects cannot execute because they are inactive disembodied signatures. Therefore, a virus alert based on a ghost positive would be substantially false.[297]

Generic virus detectors issue alerts when they detect viral evidence in a program.[298] Activity monitors and heuristic detectors monitor a network for suspicious activities, while integrity checkers look out for unauthorized changes to files.[299] These detectors frequently issue false alerts because their decision rules tend to be ambiguous. A file can change for "legitimate" reasons unrelated to

---

293. *See* SLADE, *supra* note 178, at 215.

294. *See* Andreas Marx, *Anti-Virus vs. Anti-Virus: False Positives in AV Software*, VIRUS BULL., Oct. 2003, at 17-18, *available at* http://www.av-test.org/down/papers/2003-10_vb_falsepos.pdf

295. *Id.*

296. This false positive problem can be avoided if producers of anti-virus scanners properly encrypted all their virus signatures in all parts of the program being scanned, the scanning engine, and in the virus definition files. *See id.* at 18.

297. Andreas Marx reports the following ghost positive incident. The anti-virus software, *AntiVir*, was written to disinfect systems infected with the worm Win32/Qaz. *Id.* Its disinfection routine included storing the strings "StartIE" and "qazwsx.hsq" in plain text to delete keys created by Win32/Qaz. *Id.* The presence of these strings was detected by another anti-virus product, namely Network Associates' *VirusScan*, which flagged *AntiVir* as a possible variant of the Win32/Qaz worm. *Id.*; *see also* HARLEY ET AL., *supra* note 138, at 57, 78.

298. HARLEY ET AL., *supra* note 138, at 153.

299. *Id.* at 153-58.

virus infection, resulting in a false alert by an integrity checker.[300] A heuristic detector can set its detection threshold too low and allow innocent code to trigger an alert. An alert generated in this way would, likewise, be false.

In conclusion, the truth/falsity issue of a virus alert must be resolved in the context of the technology that generated it, namely the detection technology that issued the alert and the digital properties of the detected object.

D. Verifiability

The previous Part discussed forensic verification of the truth of a virus alert. In the illustrative cases, verification was determinate: the virus alert was demonstrably either true or false. This will not always be the case. The computational logic of malicious code and the mathematical algorithm controlling its operation may be such that (1) a virus detector may classify code as viral due to the presence of, perhaps, a virus signature, yet (2) forensic analysis of the code may show that executability of the infection module is indeterminate. As a result, truth or falsity of the alert is also indeterminate. The following stylized program illustrates this phenomenon and its legal implications.

Consider a virus that has inserted its code at the end of a host program, as shown in the illustration below. This is a so-called "appending virus."[301] After appending itself to the host, the virus code inserts an algorithm at the beginning of the host. The purpose of the algorithm is to transfer control to the virus when a pre-specified mathematical condition is satisfied. Another jump routine at the end of the viral code returns control to the host program.

When the computer attempts to execute the host, the algorithm runs first. The algorithm generates a random even number greater than two and tests whether it can be written as the sum of two prime numbers.[302] If it cannot be written as the sum of two primes, control is passed to the virus. The algorithm also passes the even number it has generated to the virus code. The virus code verifies that the even

---

300. *See* FITES ET AL., *supra* note 144, at 125.

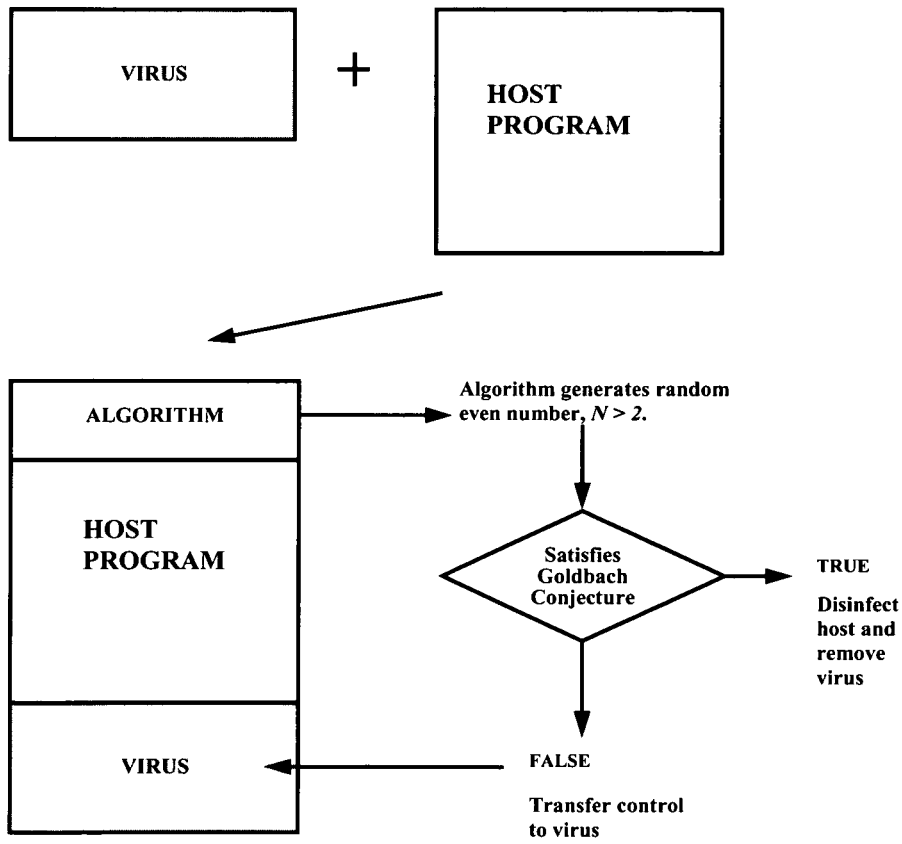
301. SKOUDIS & ZELTSER, *supra* note 137, at 36.

302. *See* GEORGE E. ANDREWS, NUMBER THEORY 15 (1994) ("A positive integer p other than 1 is said to be a prime if its only positive divisors are *1* and p."); *id.* at 20 ("If a and b ($b \neq 0$) are integers, we say b *divides* a, or b *is a divisor of* a, if a/b is an integer."). A computational method, known as Eratosthenes' sieve method, can be used to test whether an integer greater than two can be written as the sum of two primes. *See* A. Granville et al., *Checking the Goldbach Conjecture on a Vector Computer*, *in* NUMBER THEORY AND APPLICATIONS 423, 423-33 (R.A. Mollin ed., 1989).

number cannot be written as the sum of two primes and then executes, replicates, and attaches copies of itself to other host files.[303] It also executes a payload, which deletes the host computer's hard disk.

If, on the other hand, the algorithm is satisfied that the even number it has generated can be written as the sum of two primes, it does not pass control to the virus code, but removes the virus from the host program, passes control back to the host program, and modifies itself to directly transfer control to the host from then on. In other words, the virus is effectively removed from the host.

Illustration: An Indeterminate Virus



303. A virus' infection mechanism may also have a trigger. *See* HARLEY ET AL., *supra* note 138, at 7 ("[I]f the virus is at all selective about the circumstances under which it will attempt to infect, the infection mechanism may also be said to incorporate a trigger.").

Suppose this virus is detected before the host is invoked, the algorithm is executed, or the virus is triggered. It is detected, perhaps by a signature scanner—which observed the signature embedded in its main body—or by a heuristic scanner—which recognized virus-like behavior inherent in the code—such as evidence of an infection module. The vendor of the allegedly infected software believes that its quality control and business practices have been called into question by the virus alert, and sues the vendor of the antivirus software for defamation. The plaintiff needs to prove the falsity of the alert, namely that the detected object cannot execute its infection module. The defendant will prevail if she can prove the substantial truth of the alert.

A virus alert would be substantially true if the algorithm could generate a number capable of triggering the infection module.[304] This would be the case only if there exists such a number, namely an even number greater than two that cannot be written as the sum of two primes. Conversely, the virus alert would be false if such a number does not exist.

The truth or falsity of the virus alert depends on a mathematical conjecture, known as Goldbach's Conjecture.[305] The Goldbach Conjecture states that every even number greater than two can be written as the sum of two primes. For example, the even number thirty-six can be written as *17 + 19*. If Goldbach is correct, then the algorithm cannot possibly generate the kind of number that would allow execution of the viral code and its infection module. In this case, the gist of the virus alert would be provably false, and thus, the falsity issue would be decided in favor of the plaintiff.

Suppose, on the other hand, that Goldbach is incorrect. This means that there must exist at least one even number that cannot be written as the sum of two primes. If the algorithm fortuitously generated this number, its logic would transfer control to the virus, which would then execute, replicate, spread, and fire its payload. Although this number (and others that violate Goldbach, if they exist) may not be generated by the one-time run of the algorithm, the number(s) could, theoretically, be generated. The gist of the virus

304. *See* BISHOP, *supra* note 246, at 4 ("A *threat* is a potential violation of security. The violation need not actually occur for there to be a threat. The fact that the violation *might* occur means that those actions that could cause it to occur must be guarded against (or prepared for).").

305. *See* ANDREWS, *supra* note 302, at 111. The fame of the Goldbach Conjecture has even inspired a novel. *See* APOSTOLOS DOXIADIS, UNCLE PETROS AND GOLDBACH'S CONJECTURE (2000).

alert would be true in this case, and the defendant would have a constitutional defense.

The truth or falsity of an alert appears to be verifiable with mathematical precision. If Goldbach were correct, the gist of the virus alert would be false, and if Goldbach were incorrect, the gist would be true. However, Goldbach's Conjecture is an unresolved problem in mathematics. No one has (at the time of this writing) proven its truth or falsity.[306] Therefore, the truth of the virus alert is unverifiable. Under the constitutional standard articulated by the Supreme Court in *Milkovich v. Lorain Journal Co.*,[307] statements that are not objectively verifiable, or that do not contain a provably false connotation, are entitled to full First Amendment protection.[308] The virus alert in this example is therefore First Amendment protected speech, and a defamation claim based on it should be resolved in favor of the defendant. Prior to 1964 and the constitutionalization of defamation law, when the defendant had the burden of proving truth, the plaintiff would have prevailed on this issue.

VI. CONCLUSION

This article has analyzed the balance between two conflicting legal rights associated with a virus alert: the rights and guarantees of the First Amendment, and the social values protected by the law of defamation. The United States Supreme Court has addressed the conflict in a series of influential decisions.[309] The decisions raise two implications: (1) a plaintiff must plead a defamatory statement of fact that is objectively verifiable as true or false; and (2) a plaintiff must prove the falsity of the defamatory statement with convincing clarity, while the defendant may prove the truthfulness of the statement as a defense.[310]

---

306. *See* ARTURO SANGALLI, THE IMPORTANCE OF BEING FUZZY AND OTHER INSIGHTS FROM THE BORDER BETWEEN MATH AND COMPUTERS 80 (1998) ("[N]o one has yet found an even number that is not the sum of two primes; but nor has anyone demonstrated that such a number cannot exist, so the question is still unsettled."); *see also* JOHN DERBYSHIRE, PRIME OBSESSION: BERNHARD RIEMANN AND THE GREATEST UNSOLVED PROBLEM IN MATHEMATICS 90 (2003) ("Twenty-six decades of effort by some of the best minds on the planet have failed to prove or disprove this simple assertion . . . .").

307. 497 U.S. 1, 19 (1990).

308. *Id.* at 20 (interpreting *Philadelphia Newspapers, Inc. v. Hepps*, 475 U.S. 767 (1986) as "ensur[ing] that a statement of opinion relating to matters of public concern which does not contain a provably false factual connotation will receive full constitutional protection").

309. *See supra* Part II.

310. *See supra* Parts II.A-B, IV.A.

The main issues in these implications are truth, falsity, and verifiability. This article has argued that a virus alert is substantially true if, and only if, the detected object has an executable infection module.[311] A defamatory statement is substantially true if that aspect of the statement that threatens the reputation of the plaintiff is true. A virus alert threatens the plaintiff's reputation for secure software, and the reproductive capability of a virus is the essence of the security threat of a virus. Therefore, proof of such a capability is *sufficient* to prove the substantial truthfulness of a virus alert.

A statement is substantially true if it is factually similar to the proven truth and differs from the truth by no more than immaterial details.[312] This article has argued that the presence of an executable infection module in malevolent software is material to the average computer user because such a module creates a unique and serious type of risk to information security, a risk that malevolent code without such a module does not have.[313] Therefore, a virus alert is substantially true only if the detected object has a reproductive capability. Proof of such a capability is therefore *necessary* to prove the substantial truthfulness of a virus alert. Thus, proof of an executable infection module is *necessary and sufficient* to establish the substantial truthfulness of a virus alert.

The logic of these arguments is consistent with the doctrinal requirement that proof of the truthfulness of a defamatory allegation be as precise and specific as the allegation itself. Identifying an object without a reproductive capability as a virus materially mischaracterizes it and significantly misstates the nature as well as the degree of harm it is capable of causing. Furthermore, the specificity of a virus alert, as a warning that implies a risk that could escalate into an electronic tsunami as opposed to a localized threat, strengthens the credibility and impact of the communication. A defamer that uses such specificity to strengthen the credibility of the story must pay the price, namely being required to prove the truth with evidence as precise as the allegation itself. Proof of the truth of a virus alert should, therefore, include proof of an executable infection module.

This article further contributes a forensic analysis of the truthfulness of a virus alert based on the properties of the technologies that generated the alert.[314] This article merges

311. *See supra* Part V.B.
312. *See supra* Part V.A.
313. *See supra* notes 246, 247, 250-52 and accompanying text.
314. *See supra* Part V.C.

perspectives from constitutional law, the law of defamation, and information technology. Perhaps the most striking insight of the analysis is the role of technology in shaping the contours of the balance between conflicting legal interests. For instance, insights from theoretical computer science demonstrate that the truth of a virus alert may be unverifiable. In such a case, the alert would receive full constitutional protection under the Supreme Court's First Amendment defamation jurisprudence.

Although this article focuses on false positives issued by virus detectors, the analysis can be adapted to false positives in other contexts, such as wrongful mammogram, HIV, sobriety, polygraph, drug, or paternity tests. Similar constitutional, reputational, and technological issues would likely play a key role in these contexts. Plaintiffs have litigated false positives in drug tests and medical diagnoses, claiming negligence,[315] defamation,[316] and emotional distress,[317] but none of the reported defamation cases has considered the truth, falsity, and verifiability issues analyzed in this article. Thus, although the analysis in this article has focused on false positives in computer and information security, it could have wide-ranging implications for defamation jurisprudence beyond this context.

315. *See* Nehrenz v. Dunn, 593 So. 2d 915 (La. Ct. App 1992); Lewis v. Aluminum Co. of Am., 588 So. 2d 167 (La. Ct. App. 1991); Karen Manfield, Comment, *Imposing Liability on Drug Testing Laboratories for "False Positives": Getting Around Privity*, 64 U. CHI. L. REV. 287, 299-302 (1997); *see also* Thomas L. McGovern III, Note, *Employee Drug-Testing Legislation: Redrawing the Battlelines in the War on Drugs*, 39 STAN. L. REV. 1453, 1458, n.32 (1987) (discussing settlement of negligence action against testing laboratory by two job applicants whose applications were rejected due to false positive tests for marijuana).

316. *See* Willis v. Roche Biomedical Labs., Inc., 61 F.3d 313 (5th Cir. 1995); Houston Belt & Terminal Ry. Co. v. Wherry, 548 S.W.2d 743 (Tex. Civ. App. 1976).

317. *See, e.g.*, R.J. v. Humana of Fla., Inc., 652 So. 2d 360 (Fla. 1995). The petitioner, who was misdiagnosed as having Human Immunodeficiency Virus (HIV), filed suit against the medical facility, laboratory, and physician responsible for the misdiagnosis, claiming physical and mental anguish. *Id.* at 362. The Florida Supreme Court recognized that a negligent misdiagnosis that results in physical injuries (e.g. from unnecessary treatment) may be recoverable in tort. *Id.* at 363. The Court held that in this particular case, petitioner's alleged injuries were insufficient to satisfy the so-called "impact rule," which limits recovery for negligent infliction of emotional distress to physical injuries. *Id.* at 364.