Harvard Journal of Law & Technology
Volume 19, Number 2 Spring 2006

SPLOG! OR HOW TO STOP THE RISE OF A NEW MENACE ON THE INTERNET

I. INTRODUCTION

The first computer virus, more prank than malice, appeared in 1988, before the Internet became mainstream.[1] Since then, the openness of the Internet has spawned a battle between Internet innovators and abusers. This battle has resulted in an "arms race" between the two parties, each trying to best the other. Along the way, both federal and state governments have intervened to limit abuse of the Internet where possible.[2]

Recently, the conflict has spread to two new types of Internet media that have exploded in the last few years: weblogs ("blogs")[3] and

1. *See* Steve Shackelford, Note, *Computer-Related Crime: An International Problem in Need of an International Solution*, 27 TEX. INT'L L.J. 479, 484 (1992).

2. Computer viruses are now illegal. *See* 18 U.S.C. § 1030(a)(5) (2000). *See generally* Neal Kumar Katyal, *Criminal Law in Cyberspace*, 149 U. PA. L. REV. 1003, 1013–26 (2001). Recently, the federal government has also outlawed certain forms of e-mail spam. *See* Controlling the Assault of Non-Solicited Pornography and Marketing (CAN-SPAM) Act of 2003, 18 U.S.C.A. §§ 7701–7713 (West 2003). S*ee generally* Adam Hamel, Note, *Will the CAN-SPAM Act of 2003 Finally Put a Lid on Unsolicited E-mail?*, 39 NEW ENG. L. REV. 961, 979–93 (2005).

3. Blogs are websites featuring a running commentary by the blog author, or blogger, often updated multiple times per day on a particular topic. An important feature of many blogs

wikis.[4] As these new technologies have grown, so has the incidence of abuse. Since 2003, new forms of spam have been attacking the blogosphere[5] and wiki websites. This spam goes by many names, most often "link spam" and "spam blogs." Both forms of spam take advantage of the open, collaborative nature of these technologies to clog them with links and information undesired by most users of the websites. While they may seem innocuous, such forms of spam congest blogs, disrupt attempts to search for information on blogs and wikis, and dilute the information content of affected websites. Spam blogs may be "the biggest problem on the Net right now after identity theft," at least according to Mark Cuban, a well-known and outspoken Internet entrepreneur.[6] VeriSign, a leading Internet security firm, recently noted on its own blog that "the number [of spam blogs] is growing faster than the number of real blogs. By a good margin."[7] According to one report, an average of forty-four of the top one hundred results on blog search engines are now spam blogs.[8] This Note examines potential legal solutions to the proliferation of link spam and spam blogs, arguing that even though self-help by Internet innovators may mitigate this problem, regulation nonetheless is warranted.

Part II discusses how link spam and spam blogs work and the harm they cause. Part III examines the inadequacy of attempted private, technological solutions. Part IV then discusses the relevant law and suggests some potential legal solutions to the problem.

---

is that readers can post comments in response to posts and leave Internet hyperlinks to other sites of interest. *See* Wikipedia, *Blogs*, http://en.wikipedia.org/wiki/Blog (as of Mar. 15, 2006, 19:00 GMT). Seventy thousand new blogs appear each day. *See* David Kesmodel, *'Splogs' Roil Web, and Some Blame Google*, WALL ST. J., Oct. 19, 2005, at B1.

4. A wiki "is a type of website that allows users to add and edit content easily and is especially suited for collaborative writing." Wikipedia, *Wiki*, http://en.wikipedia.org/wiki/Wiki (as of Mar. 15, 2006, 22:41 GMT). The most popular of these sites is Wikipedia, a free online encyclopedia, written entirely by volunteers. Wikipedia, http://www.wikipedia.org (last visited Apr. 29, 2006). These sites present an enormous opportunity for collaborative peer production and learning. *See* Yochai Benkler, *Coase's Penguin, or, Linux and the Nature of the Firm*, 112 YALE L.J. 369, 386–87 (2002).

5. "Blogosphere" refers to the collection of all blog websites on the Internet.

6. Kesmodel, *supra* note 3.

7. Michael Graves, Welcome to the Infrablog: Weblogs 2.0, http://www.infrablog.verisign.com/2005/10/weblogs_20_1.html (last visited Apr. 29, 2006).

8. UMBRIA, INC., SPAM IN THE BLOGOSPHERE 1, 4–5 (2006), *available at* http://www.umbrialistens.com/files/uploads/umbria_splog.pdf.

II. THE NEW SPAM MENACE

A. Description of the Problem

Link spamming began in 2003,[9] and has been causing problems on the Internet ever since. Link spamming refers to the practice of repeatedly placing Internet hyperlinks on easily edited websites in hopes of increasing the number of visitors to the spammer's site.[10] Link spam works to increase traffic on a spammer's website in two ways. First, link spam is often disguised or phrased deceptively in comments to blogs. Unwitting blog readers, expecting more information, follow the link spam to the spammer's website. Second, link spam can improve the rank of a website in Internet search engine results. Most search engines, including the popular Google, rank the results of searches by the number of links "pointing" to a specific site.[11] By placing links all over the Internet in blogs and wikis, the spammer can increase his site's position in Internet search results and generate even more visitors to his site.[12]

Link spammers rely on the quantity of links, so they often use automated programs to post messages on thousands of websites and blogs at the same time.[13] They write software for this purpose and take advantage of weaknesses in Internet infrastructure to clog Internet blogs and wikis. The owners and operators of the affected sites can always remove the material, and they often do,[14] but keeping up with a machine intent on covering a site with link spam is not an easy task.

Spam blogging is a related, and even more troubling, problem that operates in much the same way. Because of the complexity involved in creating a blog from scratch, many bloggers use blog-hosting websites that feature simple, and often free, user interfaces.[15]

---

9. Wikipedia, *Spam in Blogs*, http://en.wikipedia.org/wiki/Blog_spam (as of Mar. 15, 2006, 13:11 GMT).

10. *See id*.

11. *See, e.g.*, Taher H. Haveliwala, *Topic Sensitive PageRank*, WWW2002 (2002), http://www2002.org/CDROM/refereed/127. A hyperlink "points" to a website if clicking on the link takes the user to the website. Google refers to these "pointers" as incoming links.

12. Link spammers could simply create their own sites with links to increase their ranking. But many search engines look not only at the number of links point to a site, but also the ranking of the site containing the links. Therefore, link spammers benefit by placing links to their sites on already popular websites.

13. *See* Charles Arthur, *Interview with a Link Spammer*, THE REGISTER, Jan. 31, 2005, http://www.theregister.co.uk/2005/01/31/link_spamer_interview/.

14. Alternatively, Google's blog-hosting site, Blogger.com, now allows bloggers to approve all comments before their publication on the blog. *See* Blogger Help, How Do I Moderate Comments on My Blog?, http://help.blogger.com/bin/answer.py?answer=1220 (last visited Apr. 29, 2006). Other websites allow bloggers to remove offending comments.

15. Examples of blog-hosting websites are Blogger, http://www.blogger.com (last visited Apr. 29, 2006), Yahoo!360, http://360.yahoo.com (last visited Apr. 29, 2006), and LiveJournal, http://www.livejournal.com (last visited Apr. 29, 2006).

Spam blogs, or splogs, take advantage of this infrastructure. Software, written by spammers, dumps thousands of messages at a time on blog-hosting websites.[16] The spam blogs are often gibberish or nonsensical, or they contain postings taken from other blogs.[17] Keywords are mixed into these fake blogs to attract search engines and links to the spammer's website. Some spam blogs contain advertisements, allowing the spammer to generate revenue when visitors stumble upon the spam blog from a blog search engine.[18] The links also increase the spam blog's ranking on general search engines.[19]

Both problems — link spamming and spam blogging — have been getting worse. In October 2005, spam bloggers inundated a single blog-hosting site with over 13,000 spam blogs in one weekend.[20] One Internet company estimates that 2%–8% of the 70,000 new blogs created daily are spam blogs,[21] and others place that percentage even higher.[22]

B. The Costs of Link Spam and Spam Blogs

Link spam and spam blogs inflict serious harm upon blogs and wikis. Blogs have a strong and growing following on the web.[23] They are beginning to affect society's views on technology, news, and politics.[24] Many blogs offer an alternative to mainstream media, representing an unfiltered, first-person account or an opinion that readers find trustworthy.[25] In addition, blogs have become an important form of participatory mass media, increasing the choices and information available to all.[26] Similarly, wikis and other collaborative efforts on the Internet are becoming popular and useful. Wikipedia has emerged

16. *See* Yuki Noguchi, *A New Place for Spam's Same Old Pitches*, WASH. POST, Nov. 4, 2005, at D1.

17. *See* Kesmodel, *supra* note 3.

18. *See* Noguchi, *supra* note 16 (discussing how link spammers and spam bloggers make money from advertisements).

19. *See* Wikipedia, *Spam Blog*, http://en.wikipedia.org/wiki/Spam_blog (as of Feb. 28, 2006, 16:20 GMT).

20. *See* Kesmodel, *supra* note 3.

21. *See id.*

22. *See* Graves, *supra* note 7.

23. *See* DAVID KLINE & DAN BURSTEIN, BLOG! HOW THE NEWEST MEDIA REVOLUTION IS CHANGING POLITICS, BUSINESS, AND CULTURE 5–6 (2005) (discussing the recent rise in blog readership).

24. *See id*. at 3–24 (discussing, with anecdotes, the influence that blogs have had on American politics); K. Daniel Glover, *The Rise of Blogs*, NAT'L J., Jan. 21, 2006, at 30 ("Blogs have had a noticeable impact on American society since at least 2001.").

25. *See* KLINE & BURSTEIN, *supra* note 23, at 6–8.

26. *See, e.g.*, Rebecca MacKinnon, *The World-Wide Conversation: Online Participatory Media and International News* 1, 40–50 (2004), http://media-cyber.law.harvard.edu/blogs/gems/techjournalism/WORLDWIDECONVERSATION.pdf (discussing the effects and benefits of online participatory media).

as a popular source for information on subjects from aardvark to Zanzibar.[27]

As spam blogs and link spam spread, Internet users find it more difficult to access these new forms of shared information. The surrounding noise created by web spammers hampers these new modes of collaboration and communication. Link spam may seem to annoy or frustrate only end users, but closer examination reveals that it can hinder the development of these collaborative technologies as well.[28]

If the information posted on a blog or wiki page is defaced by advertising or links, users will no longer trust that source to provide reliable information. Users may also quit using the page out of frustration when forced to sift through ads, links, and long lists of products for sale to find the rare nugget of information.[29] The collaborative nature of the site should work to correct spam, but the fight may prove too costly. For example, in 2005, the L.A. Times decided to experiment with a "wikitorial" on their website.[30] The newspaper initially posted a short editorial and invited readers to modify it. While the vast majority of participants edited the material earnestly, a few spammers inserted obscene material. The L.A. Times diligently took down the offending material, but it was consistently reposted within seconds. Two days after posting the wikitorial, the newspaper decided to take it offline.[31] This wikitorial's demise was due to obscene material, but the same decision to eliminate a wiki could easily be made following a flood of ads and link spam.

III. RESPONSE FROM THE PRIVATE SECTOR

Those afflicted by spam blogs and link spam have not sat idly by. Rather, as in other classic cases of Internet abuse, they have tried in various ways to eliminate this threat.[32] Individuals' attempts have not

---

27. *See* Benkler, *supra* note 4, at 386 (noting that Wikipedia fares no better or worse in terms of accuracy than the online version of the Columbia Encyclopedia).

28. *See* Noguchi, *supra* note 16 (quoting the editor of About.com: "[Link spam] hampers the open conversation that is the very nature of blogs.").

29. Similarly, e-mail spam inconveniences consumers and costs Internet service providers money. *Cf.* Amy G. Marino, Comment, *Is Spam the Rock of Sisyphus? Whether the CAN-SPAM Act and Its Global Counterparts Will Delete Your Email,* 32 PEPP. L. REV. 1021, 1021 (2005) (noting that, at the time CAN-SPAM was passed, 56 percent of all e-mail was spam); Lily Zhang, Note, *The CAN-SPAM Act: An Insufficient Response to the Growing Spam Problem*, 20 BERKELEY TECH. L.J. 301, 305–06 (2005) (discussing the costs of spam and the use of filters to fight it).

30. Alicia C. Shepard, *Postings of Obscene Photos End Free-Form Editorial Experiment*, N.Y. TIMES, June 21, 2005, at C8.

31. *Id*.

32. For example, copyright holders have sought to prevent illegal copying with digital rights management technology. *See* Dan L. Burk, *Legal and Technical Standards in Digital Rights Management Technology*, 74 FORDHAM L. REV. 537, 538–39 (2005). Similarly, e-mail users and providers routinely employ spam filters.

been futile, but they alone cannot adequately control spam blogs and link spam.

Private responses to web spam generally take one of three forms. First, some bloggers and blog-hosting sites respond to link spam on a case-by-case basis. When a blogger notices link spam in the comments to his or her blog, the blogger can remove the offending material from the site. Similarly, as blog-hosting websites become aware of spam blogs created in their systems, they can remove the blog from the Internet. To this end, a number of websites have been created to aid in reporting spam blogs. SplogReporter,[33] for example, encourages visitors to report spam blogs found on the Internet, with hopes of compiling a database of offending spam blogs.[34] These approaches are inadequate because machine-generated spam blogs and link spam proliferate too fast for a direct attack removal method to work.

Second, a few blog-hosting sites instead try to prevent automated programs from creating spam blogs so easily. For example, they may require a user to re-type a word printed on the screen in a distorted fashion.[35] The system is designed to allow humans to read the word but frustrates machines and automated software. Unfortunately, this method is not completely effective.[36] Some sites require a user to respond to an e-mail before signing up for a blog, but automated spamming software can defeat this method as well.[37] And, of course, both methods make it less convenient for actual users to create blogs.[38]

Third, private actors have sought to prevent spammers from improving their rankings on search engines. For link spam and spam blogs to be effective, search engines must see and count the links that the spammer scatters across the Internet.[39] With a few technical adjustments, bloggers and blog-hosting sites can allow visitors to follow the links on the pages while preventing search engines from counting the links by using "nofollow" HTML tags.[40] The growing use of this method lessens the incentives to spam blog. Spammers, though, might

---

33. SplogReporter, http://www.splogreporter.com (last visited Apr. 29, 2006).

34. The owner of SplogReporter admits that he is not sure what to do with this database. *See* Kesmodel, *supra* note 3. Thus, the utility of this approach is unclear.

35. *See* Kesmodel, *supra* note 3; Noguchi, *supra* note 16.

36. Jason Goldman, the product manager for Blogger.com "acknowledges the security feature isn't foolproof." Kesmodel, *supra* note 3.

37. *See* Arthur, *supra* note 13 (discussing how requiring e-mail responses makes spamming more difficult, but still possible).

38. According to Goldman, "The challenge is one of balance: to make it difficult for people to post bad script but not make it hard for our users." *See* Noguchi, *supra* note 16.

39. Spam in Blogs, *supra* note 9.

40. For example, blogs and wikis can use Javascript or intermediate web pages. *See id*. Google recently announced that it will no longer "count" links that contain "nofollow" in the code surrounding the links. Internet users will still be able to click on the links to follow them, but the appearance of the link will not help a website's ranking in a search result on Google. *See* Google Information for Webmasters, http://www.google.com/webmasters/bot.html#noindextags (last visited Apr. 29, 2006).

respond by increasing their activity to maximize exposure on the few sites that do not alter hyperlinks. This method also hurts the search rankings of legitimate sites: when a website becomes the hot topic of the day on thousands of blogs, search engines may overlook its importance.

IV. FASHIONING A LEGAL RESPONSE TO END LINK SPAM AND SPAM BLOGS

A. The Need for a Legal Response

Despite their increasing use and sophistication, technological responses alone have not stopped the growth of web spam.[41] Moreover, these responses have significant drawbacks: they cost money to research and implement, and lack of transparency may cause them to unknowingly block some desirable conduct.[42] One link spammer recently hinted at a supplemental solution: "While it's legal, it will continue."[43]

Legislators should respond to the problem of link spam and spam blogs because private solutions are imperfect. Without the ability to discriminate perfectly between spam and other types of posts, technology must rely on rough proxies to protect blogs and wikis from unwanted material. For instance, some blog-hosting websites charge a fee for their service, using willingness to pay as a proxy for legitimacy.[44] Those sites host few, if any, spam blogs, but increased costs depress the availability of legitimate as well as illegitimate blogs. Similarly, some bloggers exclude all comments from their site rather than suffer the frustration of comment spam. This result threatens the role of blogs as an interactive medium, rendering readers unable to provide relevant information, helpful links, and insightful commentary.[45] The risk of over-deterrence in private solutions enhances the desirability of a properly tailored legal solution.

Legislation can address spam blogs and link spam without harming the interests of non-abusive users. Unlike private actors, who must rely on indirect responses dependant upon code and manpower, legis-

---

41. *See* Noguchi, *supra* note 16. Also, Jason Goldman, the product manager for Blogger.com, recently stated that spam blogs would be a problem for the blogging community for some time to come. Kesmodel, *supra* note 3, at B1.

42. *See* David E. Sorkin, *Technical and Legal Approaches to Unsolicited Electronic Mail*, 35 U.S.F. L. REV. 325, 356 (2001) (discussing the limitations of technical approaches to dealing with e-mail spam).

43. Arthur, *supra* note 13.

44. Yahoo!, a popular Internet portal and search engine, offers blog-hosting services for a fee on its small business website. Yahoo! Small Business, http://smallbusiness.yahoo.com/webhosting/problogs.php (last visited Apr. 29, 2006).

45. *See* MacKinnon, *supra* note 26, at 19–21 (discussing the importance of the comment section of a blog).

lators can tailor statutory language to directly proscribe abuses.[46] Of course, any law's effectiveness depends on its enforcement: an anti-spam law cannot succeed unless government entities either supply the resources needed to enforce it or provide individuals with an incentive to police violations.[47] External legal limitations — most notably, the First Amendment — may present an additional obstacle.[48] Nonetheless, the legislature should not wait out the arms race; it should act now to supplement the necessarily inadequate private responses to spam and to avoid overzealous private enforcement efforts.

But will it work? Critics might argue that legal solutions, unlike technological ones, cannot keep up with emerging Internet technologies. For instance, Congress's attempt to control e-mail spam has yielded mixed results.[49] Similarly, despite the long-standing protection of copyright laws, recording companies are turning to technology to prevent illegal digital copying.[50] Furthermore, an effective law may push spammers overseas, out of reach of government enforcement.[51] These arguments erroneously imagine a legislative response supplanting a technological one; Congress should pass legislation in this area only to *complement* the technological responses.[52] Legislation can fill

---

46. Furthermore, it may be desirable to choose the response to spam collectively, or through the political process, rather than through private action. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 219–20 (1999); Jay P. Kesan & Rajiv C. Shah, *Shaping Code*, 18 HARV. J.L. & TECH. 319, 325 & n.27 (2005) (describing the position of those who say "the government can and should" shape the Internet with regulation).

47. The magnitude of enforcement costs will depend on the structure of the legal solution and may be low in comparison to the costs associated with other regulatory regimes. *See* Kesan & Shah, *supra* note 46, at 329 (arguing that prohibitions often have lower enforcement costs than other forms of regulation).

48. *See infra* Part IV.C.

49. *Compare America Online Reports Drop in Spam*, N.Y. TIMES, Dec. 28, 2004, at C5 (citing AOL's report of a large decrease in the amount of spam its users received, a year after CAN-SPAM was passed), *with* Jonathan Krim, *Senate Hears Mixed Reviews of Anti-Spam Law; Some Say Consumers Need More Protection*, WASH. POST, May 21, 2004, at E5 (describing the congressional testimony of many that spam had not declined in the six months after CAN-SPAM was passed), *and* David McGuire, *A Year After Legislation, Spam Still Widespread; Technology Seen as Best Deterrent*, WASH. POST, Jan. 4, 2005, at E5 (noting no drop-off in spam a year after CAN-SPAM was passed).

50. *See* Julie E. Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J. 575, 580–88 (2003) (discussing digital rights management and how it can protect intellectual property rights).

51. *Cf.* Sorkin, *supra* note 42, at 356 (describing enforcement problems caused by jurisdictional limitations for e-mail anti-spam laws).

52. The experience of legislative efforts to curb e-mail spam and copyright infringement does not suggest that legislation cannot play an important role in limiting spam blogs and link spam. In the context of e-mail spam, the CAN-SPAM Act was designed to provide choice for e-mail users, not to slow the rise of spam: it enabled individual opt-out while leaving spammers free to spam. *See* Zhang, *supra* note 29, at 317–19 & 326 (noting that CAN-SPAM could lead to more spam because it legitimizes some spam). Thus, statistics noting a rise in the number of spam messages sent do not indicate that the legislation necessarily failed. In the context of copyright infringement, regulators faced an uphill battle because they sought to prevent diffuse violations: because of the massive number of infringers, each infringer faced a low chance of detection. There are undoubtedly fewer spam bloggers than MP3 downloaders, arguably permitting easier enforcement.

in the shortcomings of a purely private response and help guard against an overzealous response, promoting development of an important arena for free speech and collaboration.[53]

B. Three Possible Legal Approaches

Within the wide range of potential legal solutions to web spam, this Recent Development focuses on three popular and intuitive options.

1. Ban Undesirable Content

The most obvious and complete legislative response is to proscribe all spam blogs and link spam. By banning the conduct altogether, the law could eliminate undesirable and unwanted clutter without any direct harm to the general development of blogs or wikis. A reasonable level of enforcement also would make a proscription of this kind fully effective, since it would reach all Internet spam. At minimum, a proscription statute would consist of two elements: a definition of link spam and spam blogs and a statement of either civil or criminal penalties for those who post spam blogs or links.[54]

2. Ban the Best Methods of Creating Undesirable Content

Rather than proscribing the *content* of spam blogs and links, legislation could proscribe certain *methods* of producing blogs and comments. Automated software makes spam blogs possible because it enables a spammer to create thousands of blogs per day.[55] A ban on the use of such software to create blogs or comments would force spammers to enter their advertising blogs manually, leaving them unable (because of limited manpower) or unwilling (because of in-

---

53. Given the nationwide (and worldwide) use and appeal of the Internet, Congress, rather than individual states, should make these laws.

54. The question of civil versus criminal enforcement is orthogonal to the question of what method of regulation to employ. While a purely civil liability regime would leave enforcement in the hands of private individuals with a direct incentive to eliminate the problem, it might fail for three reasons. First, web spam produces diffuse harms — to readers as well as bloggers — such that no one potential plaintiff internalizes a large share of its costs. Second, bloggers as plaintiffs might lack the resources necessary to identify the worst offenders and to fund discovery necessary to build a strong case against them. Third, the old doctrine of trespass to chattels arguably imposes civil liability on web spammers but has not curbed web spam. *See* eBay v. Bidder's Edge, 100 F. Supp. 2d 1058 (N.D. Cal. 2000); Daniel Kearney, Note, *Network Effects and the Emerging Doctrine of Cybertrespass*, 23 YALE L. & POL'Y REV. 313, 318–23 (2005) (discussing generally the tort of cybertrespass and the *eBay v. Bidder's Edge*). Consequently, a criminal liability regime, or a regime of joint civil and criminal liability, probably would permit more effective enforcement.

55. *See* Arthur, *supra* note 13; Noguchi, *supra* note 16.

creased operating costs) to create as many blogs.[56] Generally, spammers "look[] to spam not for quality [of site] but quantity of links."[57] By eliminating their ability to create a massive web of links cost-effectively, the law would attenuate their impact on search results and therefore their incentive to spam. Blog-hosting sites, such as Google's Blogger.com, have implemented technologies designed to prevent the use of automated software, although these technologies have not been completely effective.[58]

A law regulating methods of creating spam blogs and link spam would define blogs, blog-hosting sites, and wikis. It would then proscribe the use of scripts, bots, and other programs to quickly post multiple messages on these sites.

3. Force Producers to Label Undesirable Content

Borrowing from the approach of the CAN-SPAM Act,[59] Congress could require all spam blogs and link spam to include a stock message identifying them as spam. This labeling requirement could apply to all link spam and spam blogs or only to messages created by automated programs. By requiring a set phrase — such as "This message was created by an automated program" — it would allow Internet users to recognize spam messages easily. More importantly, blog-hosting sites and bloggers could identify and remove offending material quickly by programming software to eliminate all messages containing the required tagline.

C. Evaluating How to Proceed

Three factors affect the desirability of any proposal. First, practicality: Can legislators draft statutory language that will effectuate the proposed solution? If so, can the solution be enforced? Second, effectiveness: Would the proposed solution eliminate or reduce the prevalence of spam blogs and link spam without unduly harming blogs and

---

56. The law could also proscribe the use of the open proxies necessary to relay the unwanted spam. Open proxies are computers that can be accessed by outside users that relay them to websites and web services. Wikipedia, *Open Proxy*, http://en.wikipedia.org/wiki/Open_proxy (as of April 13, 2006, 19:55 GMT). Spammers often use these open computers to relay their spam and disguise its true source. *See* Arthur, *supra* note 13.

57. Arthur, *supra* note 13.

58. Despite the introduction of screening technology, the number of spam blogs continues to rise. *See supra* note 22 and accompanying text. Also, automated software often can beat these technologies. *See supra* note 42 and accompanying text; Michael Pollit, *Cashing in on Fake Blogs*, THE GUARDIAN, Nov. 17, 2005, *available at* http://technology.guardian.co.uk/weekly/story/0,16376,1643774,00.html (arguing that the battle against spam bloggers is futile because spam bloggers learn quickly how to beat anti-spam measures).

59. 18 U.S.C.A. § 7704(d) (2000) (requiring warning labels on sexually explicit commercial e-mails).

wikis? Third, constitutionality: Does the proposal stay within First Amendment limits in its restriction of the speech of spam bloggers and other Internet users?

Because of the complexity of First Amendment doctrines regarding advertising and "nuisance" speech, the issue of constitutionality presents the most difficult problem. Spam blogs and link spam, while often unintelligible, probably constitute speech and therefore receive some form of protection from the First Amendment.[60] Yet the First Amendment allows considerable room for the regulation of activity contained in the category of "speech." Specifically, it permits significant limitations on commercial speech, as well as restrictions on the time, place, and manner of both commercial and non-commercial speech.

The Supreme Court has often stated that commercial speech receives, if not less protection, *different* protection than traditional non-commercial speech.[61] The federal government utilized this distinction in crafting legislation to combat e-mail spam in 2003.[62] If spam blogs and link spam are commercial speech, courts would judge any law regulating them under the test enunciated in *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York*.[63] Under this test, Congress must have a substantial interest in regulating the speech, and there must be a reasonable fit between the regulation and the substantial interest.[64] In other words, any regulation on commercial spam blogs and link spam must "directly advance[] the governmental interest" in protecting bloggers and blog readers from annoyance, harassment, and noise in a manner "not more extensive than is necessary to serve that interest."[65]

For non-commercial speech, courts apply strict scrutiny to content-based laws and typically strike them down.[66] Nevertheless, Congress can enact appropriate time, place, and manner restrictions on speech without violating the Constitution.[67] Thus, the First Amendment analysis turns on whether spam blogs are commercial speech or not and whether the regulation is a (presumptively unconstitutional)

---

60. *See* Timothy Wu, *Application-Centered Internet Analysis*, 85 VA. L. REV. 1163, 1170 (1999) (concluding that the Supreme Court's decision in *Reno v. ACLU*, 521 U.S. 844 (1997), amounts to a rule entitled "The Internet Gets Full First Amendment Protection").

61. *See* Cincinnati v. Discovery Networks, Inc., 507 U.S. 410, 422 (1993); Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447, 455–56 (1978).

62. *See* Zhang, *supra* note 29, at 318 ("Although the Act targets spam, it is solely applicable to spam messages which are commercial in nature.").

63. 447 U.S. 557, 566 (1980).

64. *Id*.

65. *Id.* at 566. According to the Supreme Court, the last portion of the test does not impose a least-restrictive-means requirement. *See* Bd. of Tr. of the State Univ. of N.Y. v. Fox, 492 U.S. 469, 477–80 (1989).

66. *See* Police Dept. of Chi. v. Mosley, 408 U.S. 92, 99 (1972).

67. *See* Ward v. Rock Against Racism, 491 U.S. 781, 791–92 (1989); Clark v. Cmty. for Creative Non-Violence, 468 U.S. 288, 293 (1984).

content-based restriction or a (presumptively constitutional) time, place, or manner restriction.

1. Is Spam Commercial Speech?

The Supreme Court has drawn the line between commercial and non-commercial speech in a number of ways, leaving the "precise bounds" of the category "subject to doubt."[68] At its core, commercial speech consists of "speech which does no more than propose a commercial transaction."[69] This core definition covers some link spam and spam blogs. For instance, spammers often post long lists of products with prices as comments to blogs. Similarly, spam blogs often contain cryptic offerings of low home mortgage rates or low priced goods.[70] These statements, of the form "I will sell you the X . . . at the Y price," are undoubtedly core commercial speech.[71] Although advertisements may contain more than a simple offer of a good for a price, "advertising pure and simple" also fits within this category.[72]

Other commercial spam blogs, however, do not propose a transaction or advertise a product at all, seeking only to increase traffic to the spammer's website. Had the Court restricted its definition of commercial speech to core commercial speech, it might have excluded such indirect advertising. Yet the Court has eschewed such formalism and adopted a more expansive definition.[73] In its broadest formulation, the Court included any "expression related solely to the economic interests of the speaker and its audience."[74] Since businesses use spam

68. Zauderer v. Office of Disciplinary Counsel of Supreme Court of Ohio, 471 U.S. 626, 637 (1985); *see also* Robert Post, *The Constitutional Status of Commercial Speech*, 48 UCLA L. REV. 1, 5–6 (2000) (discussing what encompasses commercial speech and stating that "[t]he boundaries of the category are thus quite blurred").

69. Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 66 (1983) (quoting Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 762 (1976) (quoting Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 385 (1973))) (internal quotation marks omitted).

70. For instance, the Wall Street Journal recently quoted one spam blog as reading, "Cool blog. I have a home equity loan lowest rate blog myself. It's [sic] goes over home equity loan lowest rate. Please visit, thanks!" Kesmodel, *supra* note 3, at B1. The spam blog contained a link to a web site offering home loans. *Id.*

71. Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 762 (1976).

72. *Zauderer*, 471 U.S. at 637.

73. In deciding whether contraceptive pamphlets constituted commercial speech, the Court balanced three factors: whether the pamphlets were advertisements, whether they referenced a specific product, and whether the sender had an economic motive for sending them. *Bolger,* 463 U.S. at 66–67. No one factor was determinative. *Id*. at 67. The Ninth Circuit has questioned whether these factors can still be used to define non-core commercial speech. Thomas v. Anchorage Equal Rights Comm'n, 165 F.3d 692, 710 (9th Cir. 1999), *rev'd en banc on other grounds*, 220 F.3d 1134 (9th Cir. 2000).

74. Cent. Hudson v. Pub. Serv. Comm'n, 447 U.S. 557, 561 (1980). Since then, it has backed off from that definition, requiring a more narrow examination "to ensure that speech deserving of greater constitutional protection is not inadvertently suppressed." Cincinnati v.

only to increase sales of their products, albeit indirectly, a court probably would treat all web spam tied to a commercial website as commercial speech.

Yet non-commercial entities also may have cause to use spam blogs and link spam: political groups and ideological organizations often seek to increase exposure by increasing traffic to their websites, just as commercial businesses do. While non-commercial spam blogging is uncommon at the moment, the possibility of its existence renders any law subject to a First Amendment overbreadth challenge if it (a) targets all spam blogs and (b) satisfies First Amendment restrictions only because of the Court's differential treatment of commercial speech.[75] To take advantage of the relatively weak protection afforded to commercial speech, then, legislators must draft a narrow statute, targeting only commercial link spam or spam blogs made to profit commercially.[76]

2. Misleading or Deceptive Spam

One of the most significant consequences of the distinction between commercial and non-commercial speech concerns the treatment of misleading or deceptive speech: Congress can ban misleading or deceptive *commercial* speech, as it falls outside of First Amendment protection.[77] In one regard, all spam blogs and link spam are deceptions; they consciously try to pose as normal, human-created blogs, fooling readers and search engines alike. On this theory, a court could find all commercial spam able to be proscribed by Congress. But some link spam and spam blogs are quite upfront about the fact that they are advertisements. For instance, the link erroneously labeled to fool readers is deceptive, but the blog comment straightforwardly advertising a poker site is not. Furthermore, even the mislabeled link is not putting the reader in danger of striking a deceptive bargain; in-

---

Discovery Networks, Inc., 507 U.S. 410, 423 (1993) (citing *Bolger*, 463 U.S. at 66). Nevertheless, courts generally have employed an expansive definition in cases involving speech purely motivated by the desire to sell goods. For example, the Fifth Circuit recently held that e-mail spam is commercial speech, applying the broad *Central Hudson* definition. White Buffalo Ventures, LLC v. Univ. of Tex., 420 F.3d 366, 374 (5th Cir. 2005) (quoting *Central Hudson*, 447 U.S. at 561).

75. Even if all known spam blogs are related to commercial speech, the fact that spam blogs are not by necessity related to commercial websites could leave any law targeting them open to an overbreadth challenge. In the special context of the First Amendment, even if the government only enforced a regulation on spam blogging against commercial spammers, the spammers would be able to mount a defense claiming that the law is substantially overbroad. *See* Virginia v. Hicks, 539 U.S. 113, 118 (2003); Bd. of Tr. of the State Univ. of N.Y. v. Fox, 492 U.S. 469, 481–82 (1989).

76. *Cf.* CAN-SPAM Act of 2003, 18 U.S.C.A. §§ 7702, 7704 (West 2003) (defining commercial electronic messages and limiting portions of the act to commercial electronic messages).

77. *Bolger*, 463 U.S. at 69; *Fox*, 492 U.S. at 475; *Central Hudson*, 447 U.S. at 566.

stead, it merely offers the reader information she did not expect and may not want. Nonetheless, a court should find it to be deceptive within the meaning of the case law for the following reason: a contrary position leads to the absurd result that an advertisement is able to be regulated, but as it gets more confusing and incorrect, it gains First Amendment protection. Thus, for speech falling within this class of behavior, First Amendment challenges will fail.[78]

D. A Call to Action

Evaluating the potential ways Congress could address the ills of spam blogs and link spam along these three axes — practicality, effectiveness, and constitutionality — makes the correct course of action clear. Congress should, without delay, proscribe the use of automated programs and bots to create blogs and wikis, effectively limiting the proliferation of machine-generated blogs, wikis, and blog comments.

The most obvious solution — banning outright spam blogs and link spam — is both impractical and unconstitutional. As a practical matter, it is unclear how to draft such a ban. The law, of course, cannot simply declare spam illegal; it would have to define spam blogs and link spam, a difficult task. The web is littered with multiple varieties of spam blogs and link spam which pose as legitimate speech. One link spammer posts messages to his fake buddy Ned, with a link to a gambling site. Another spam blogger copies text from other blogs and includes links to hair product sites. Writing a narrow law that prohibits both of these *messages*, rather than the act of posting them, is no mean feat. Any such law runs the risk of proscribing non-spam messages and blogs as well as actual link spam and spam blogs. And since the whole point of a spam ban is to foster the development of a growing speech forum, this sort of broad law would be counterproductive.

Moreover, an outright ban on spam blogs and link spam would violate the First Amendment. A broadly written proscription on spam blogs and link spam presents two constitutional dangers. First, application of a ban on spam blogs to non-commercial speech is probably unconstitutional. A blanket ban on spam blogs necessarily would be content-based; to determine its applicability, government enforcers would have to evaluate the content of the blog. As such, the law

78. Unfortunately, unless a court adopts the position that all spam is inherently misleading, relying on the misleading and deceptive speech exception to First Amendment protection will not be fruitful. Spammers can change their behavior to regain First Amendment protection. For instance, rather than posting a home loans link on a political blog that says "Great Article. Click here to read more," spammers might choose to post something that says, "Low Home Loans. Click here." The nuisance remains without the attempt at deception.

would be subject to strict scrutiny.[79] Given that many other less restrictive alternatives remain,[80] such a ban almost certainly would be struck down.[81]

Even if the law applied only to commercial speech, it probably would be unconstitutional. The Supreme Court has been reluctant to enforce complete bans on even purely commercial speech. Indeed, the Court "review[s] with special care regulations that entirely suppress commercial speech in order to pursue a nonspeech-related policy."[82] In later cases, the Court "identified the serious First Amendment concerns that attend blanket advertising prohibitions that do not protect consumers from commercial harms."[83] Detractors of a ban would argue that even if all web spam is found to be commercial speech, a complete ban would be for a nonspeech-related reason — to lower Internet user frustration. In reality, the government's interest is broader than that. The government would seek to eliminate spam blogs, a form of pernicious commercial speech, to protect other blogs and wikis, a burgeoning form of electronic free speech.[84] But the Supreme Court's decision in *Cincinnati v. Discovery Network* makes clear that the government cannot pick and choose among types of speech even in this way.[85] In that case, the city of Cincinnati attempted to ban racks containing commercial handbills on the city sidewalks. The Court struck down the ban, noting that the city "ha[d] not asserted an interest in preventing commercial harms."[86] Instead, the city attempted to remove an "eyesore."[87] The logic of *Discovery Network* poses a problem for a complete ban on spam blogs and link spam: the government is attempting to address a nuisance, or "eyesore," on the Internet, caused by commercial and (potentially) noncommercial speech, by targeting commercial speech only. Therefore, a blanket ban on spam would not be constitutional, even if it only applies to commercial speech.[88]

---

79. *See* Police Dept. of Chi. v. Mosley, 408 U.S. 92, 99 (1972).

80. For a description of possible less restrictive alternatives, such as a labeling requirement and regulation of automated postings, see *supra* Part IV.B.

81. Reno v. ACLU, 521 U.S. 844, 874 (1997) (noting in an Internet speech case that "when a statute regulates the content of speech," it is "unacceptable if less restrictive alternatives would be at least as effective in achieving" the government's legitimate interest).

82. *Central Hudson*, 447 U.S. at 566 n.9.

83. 44 Liquormart, Inc. v. Rhode Island, 517 U.S. 484, 499 (1996).

84. Although this reasoning probably does make the government's interest "speech-related," that interest still does not relate to the *commercial* nature of the restricted speech.

85. 507 U.S. 410 (1993).

86. *Id*. at 436.

87. *Id*. at 425.

88. Recently, though, the Fifth Circuit interpreted the *Central Hudson* test as permitting Texas's ban on certain commercial e-mail spam. White Buffalo Ventures, LLC v. Univ. of Tex., 420 F.3d 366, 374 (5th Cir. 2005). In *White Buffalo*, the University of Texas, a state university and an arm of the state, completely blocked incoming spam from a certain sender. After determining that the e-mails were commercial speech, the Fifth Circuit had no trouble dismissing the spammer's First Amendment claims. *See id*. at 378. This case was special,

A labeling solution would fare somewhat better constitutionally and practically, but constitutional considerations still would limit its effectiveness. For commercial speech, the Supreme Court has stated on a number of occasions that required disclosures are less intrusive than proscriptions and thus are generally constitutionally acceptable.[89] In other words, the proposed labeling requirement likely would satisfy all four parts of the *Central Hudson* test.[90]

The Second Circuit, in *National Electrical Manufacturers Ass'n v. Sorrell*,[91] has gone as far as deciding that *Zauderer*, not *Central Hudson*, must control cases involving compelled, truthful commercial speech, however.[92] As the Second Circuit applied *Zauderer*, there must be "a rational connection between the purpose of a commercial disclosure requirement and the means employed to realize that purpose."[93] The Second Circuit also held that the purpose of such disclosure does not have to be protecting consumers from deception. In *National Electrical Manufacturers*, like in the case of web spam, the state was not attempting to protect consumers from bad bargains.[94] Therefore, under either *Zauderer* or *Central Hudson*, the labeling proposal is probably constitutional as applied to commercial speech.

Interestingly, however, because "purely commercial speech is more susceptible to compelled disclosure requirements"[95] than non-commercial speech, a labeling requirement must be restricted to commercial spam blogs and link spam only. In general, a court likely will apply strict scrutiny to any compelled disclosure requirement on non-commercial speech.[96] Therefore, depending on how much spam

---

however, because the state was also acting as an Internet provider, *see id*. at 371, and e-mail users had complained about the specific spam, *see id*. at 369.

89. *See* 44 Liquormart, Inc. v. Rhode Island, 517 U.S. 484, 498 (1996) ("Specifically, we explained that the State may require commercial messages to 'appear in such a form, or include such additional information, warnings, and disclaimers, as are necessary to prevent its being deceptive.'" (quoting Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 772 n.24 (1976))); Zauderer v. Office of Disciplinary Counsel of Supreme Court of Ohio, 471 U.S. 626, 651 (1985) ("Thus, in virtually all our commercial speech decisions to date, we have emphasized that because disclosure requirements trench much more narrowly on an advertiser's interests than do flat prohibitions on speech, 'warning[s] or disclaimer[s] might be appropriately required . . . in order to dissipate the possibility of consumer confusion or deception.'" (quoting *In re* R.M.J., 455 U.S. 191, 201 (1982))).

90. *See* Post, *supra* note 68, at 26–28 (concluding that compelled disclosures of commercial speech do not offend the constitution).

91. 272 F.3d 104 (2d Cir. 2001).

92. *Id*. at 115 ("*Zauderer,* not [*Central Hudson*], describes the relationship between means and ends demanded by the First Amendment in compelled commercial disclosure cases.").

93. *Id*.; *accord Zauderer*, 471 U.S. at 651.

94. *Nat'l Elec. Mfrs.*, 272 F.3d at 115 ("To be sure, the compelled disclosure at issue here was not intended to prevent 'consumer confusion or deception' per se, but rather to better inform consumers about the products they purchase." (quoting *Zauderer*, 471 U.S. at 651)).

95. Riley v. Nat'l Fed'n of the Blind of N.C., Inc., 487 U.S. 781, 796 n.9 (1988).

96. *See* Post, *supra* note 68, at 26 (discussing the Supreme Court's approach to compelled non-commercial speech). *See generally Riley*, 487 U.S. at 797.

can be rightly classified as commercial, the labeling requirement might be ineffective.

Given the problems with the previous proposals, Congress should enact a law proscribing the use of automated software to post to blogs, wikis, and blog comments. Because this approach would not target speech directly, the government can constitutionally attack the incentives of spammers.[97] First, the proscription should codify the *Central Hudson* test for commercial speech. The government has a substantial interest in protecting the "user efficiency" of bloggers and Internet readers[98] and the vitality of an important new method of speech. Also, this method of furthering the government's interest is a "reasonable fit."[99] It directly advances the government's interests by limiting the quantity of spam blogs and freeing up the blogosphere for productive free speech activity. Furthermore, it is not more extensive or intrusive then it needs to be, since it prevents spam blogs from proliferating in great numbers but does not prevent any particular type of speech from being posted to the Internet. In fact, the law would function much like certain portions of the CAN-SPAM Act, already enacted into law.[100]

A ban on automatically created spam blogs and link spam should withstand constitutional analysis even if some spam is found to be non-commercial speech. The proposed regulation is content-neutral in that it is "justified without reference to the content of the regulated speech" posted to the Internet.[101] Any currently posted spam blog could be re-posted without offending the new law, as long as it is not reposted with automated software. As such, the law is a content-neutral manner restriction on posting material to the Internet.[102] Furthermore, it is an acceptable manner restriction because it is narrowly tailored to the problem being addressed — the large quantity of spam blogs and comment spam — and "leave[s] open ample alternative channels for communication of the information."[103] As noted, the

97. Congress utilized a similar approach — attacking methods of creating spam rather than the spam itself — in drafting the CAN-SPAM Act. In a report relating to certain criminal elements of the Act, Sen. Orrin Hatch determined that the law "does not raise concerns under the First Amendment" because "rather than targeting speech, the bill instead targets e-mailing techniques used to steal computer services and trespass on private computers and computer networks." S. REP. NO. 108-170, at 4 (2003).

98. White Buffalo Ventures, LLC v. Univ. of Tex., 420 F.3d 366, 376 (5th Cir. 2005).

99. Bd. of Tr. of the State Univ. of N.Y. v. Fox, 492 U.S. 469, 480–81 (1989) (holding that the last two prongs of the *Central Hudson* test do not form a least-restrictive-alternative test, but rather a reasonable fit requirement).

100. One provision of that law prevents e-mail spammers from compiling e-mail lists in certain abusive ways. 15 U.S.C.A. § 7704(b)(1) (West 2005). Another provision prohibits a number of technical methods for creating e-mail addresses. *Id*. § 7704(b)(2). In addition, the CAN-SPAM Act prohibits individuals from "knowingly . . . relay[ing] or retransmit[ting] a commercial electronic mail message . . . from a protected computer or computer network that such person has accessed without authorization." *Id*. § 7704(b)(3).

101. Clark v. Cmty. for Creative Non-Violence, 468 U.S. 288, 293 (1984).

102. *See id*.; Ward v. Rock Against Racism, 491 U.S. 781, 791–92 (1989).

103. *Clark*, 468 U.S. at 293.

spammers can still use the same forums and avenues for spamming, just without the benefit of automated programs and open proxies. Indeed, such a regulation would be akin to laws that prevent the use of loudspeakers on city streets[104] or limit decibel levels at concerts.[105] Spammers can still get their "message" across, just at lower "volumes."

Finally, this regulation would be both practical and effective. As discussed earlier, drafting regulations to target automated programs is feasible. And enforcement, while difficult, is not impossible.[106] Once a user stumbles upon a nonsensical spam blog, he immediately suspects that it was machine-generated. Armed with that knowledge, enforcement authorities would have cause to investigate. At the very least, enforcement should be no more difficult than for e-mail spam, whose perpetrators prosecutors have managed to identify and charge.[107] Finally, a law restricting the most prevalent method of creating web spam would solve the main problem it poses, its ubiquity.

V. CONCLUSION

Spam blogs and link spam present a real problem to the growth of a new form of communication and media on the Internet. Like the much maligned e-mail spam, they present a headache to readers and users. Moreover, they also dilute and distract from open debate and could lower levels of blog and wiki readership. Therefore, Congress should adopt legislation prohibiting the use of software to post machine-generated spam blogs and link spam. Such a law passes constitutional muster and is likely to help in the fight against spam.

Detractors of the legal approach outlined in this Note may suggest that the problem is not large enough to warrant legislation. While most Internet users utilize e-mail, most still do not frequent blogs or wikis. Such a response, however, is short-sighted. Web spam may be slowing the growth of a new and important sector of the Internet. Spam blogging makes finding useful information in the blogosphere difficult and may discourage a marginal or first-time user from continuing to access this media. By adopting a legal solution to work in tandem with private responses, rather than in lieu of them, Congress can contribute to the growth of the blogosphere. At the very least, the legal solution proposed here would not interfere with technological attempts to end spam and would only be one more weapon in the arsenal of well-meaning Internet innovators.

---

104. Kovacs v. Cooper, 336 U.S. 77 (1949).

105. *Ward*, 491 U.S. at 784–87.

106. *Cf.* Karin H. Cather, *Canning Spam: The Nation's First Felony Spam Trial*, 39 PROSECUTOR 26, 26–28 (2005) (describing the evidence necessary to prosecute an e-mail spam case and noting that it is similar to other white collar crime cases).

107. *Id*. at 26.