Security and document compatibility for electronic refereeing

Andrew W. Appel
Department of Computer Science
Princeton University
appel@princeton.edu

October 9, 1996

I have recently described [1] a management system, with accompanying software, for use by journal editors in organizing the refereeing of scientific manuscripts prior to publication. For me this system has been very successful, cutting the time from submission to final acceptance of manuscripts from two or three years (under the previous editor) to just a few months. This is fast by the standards of computer science, where referees are meticulous but a bit slow, and almost every paper requires revision and re-refereeing before acceptance.

Security

I have been asked to comment on the security of e-mail refereeing: can manuscripts be stolen prior to publication?

In computer science, almost all articles submitted to journals have already appeared as technical reports on their authors' Web pages. Many submitted articles have already appeared in proceedings of conferences (these articles are usually substantially revised in response to referees' demands, and thus the journal article becomes significantly different from the conference version). Unpublished articles in computer science (and also in physics and mathematics) are distributed widely and freely.

Therefore, someone wishing to "steal" unpublished material in computer science or physics would not bother to penetrate the security of a journal editor's computer, when the authors are trying as hard as they can to give the articles away prior to publication.

In scientific disciplines that prohibit prepublication distribution of material, security is more of an issue. Also, in any discipline (including computer science) that protects the anonymity of referees, the "referee database" on the editor's computer must be protected against intruders.

Because my electronic refereeing system works entirely by e-mail, the editor can operate behind an Internet firewall [4]. E-mail to and from the outside world (authors and referees) can go safely through the firewall, while other kinds of access (such as "login" and "ftp") are prohibited. Thus, the editor's computer can be made reasonably secure, provided that the firewall itself is managed competently.

The security of transmitted documents and the anonymity of referees can be compromised by the fact that E-mail itself is insecure. E-mail messages can be watched by "packet sniffers" on any of the subnets between the sender and receiver. To reduce the scope of this problem, the editor's machine can be connected through a firewall whose "mail exchange" [7] protocol insists that messages be sent directly to it, rather than be queued by a third party. Then, assuming the Internet service providers (AT&T, MCI, etc.) are reasonably secure, the packet sniffer would have to run in the local network of the sender (referee) or the receiver (editor).

Another solution is to encrypt messages using packages such as Zimmerman's *Pretty Good Privacy* [6]; public-key encryption allows the author to send encrypted mail to the editor without prior secure communication. However, the identity of the sender and recipient will still be obvious; to preserve full anonymity, the editor and referee should communicate using encrypted mail through through an anonymous remailer [2, 3], which hides the identity of the correspondents from the outside world. Using public-key authentication, the referee and editor can be reasonably sure of each others' identity even though anonymous mail is used. It may be a few years, however, before encrypted e-mail is routinely used by nonspecialists on the Internet.

As editor of TOPLAS, I keep files on machines administered by the Department of Computer Science at Princeton University. Our technical staff administers e-mail and file servers on these machines expertly and with attention to security issues. I regard the TOPLAS files, and correspondence with referees, as moderately but not very secure.

The editor's machine may be secure, but the editor e-mails articles out to referees all over the world. The editor can reasonably request of the referee not to distribute copies the article, but it is impossible to require that the referee's machine be secured behind a competently managed firewall—this is simply beyond the control of the referee. Thus, there is a potential security problem with manuscripts on referees' machines.

A referee often forwards a manuscript to a colleague just because she lacks the time to review it. With the e-mail refereeing system, when the editor sends a paper he simply requests "if you wish to recommend a colleague to referee this manuscript, simply respond by e-mail." Because this is so easy for the referee to do in the e-mail reader, the editor usually gets a chance to approve or disapprove the suggestion before the referee actually forwards the manuscript. In fact, the editor simply e-mails the manuscript directly to the colleague, and the original referee is now "out of the loop." This is a very satisfactory situation: the editor is corresponding directly with the "real" referee, and the original referee is not bothered anymore.

Photographs

In my subdiscipline, of computer scientists who do research in programming languages and compilers, papers do not generally contain photographs. In other fields, such as computer graphics or biology, many papers contain high-resolution photographs, often color photographs. How are these to be exchanged by e-mail?

In the future, we may expect that the biologist's microphotograph may be obtained from an electronic scanning camera directly attached to the microscope (today, high-resolution scanning cameras are slow and expensive, so the computer-happy biologist takes 35mm slides and converts them to electronic form using a cheap but high-resolution scanner). The scientist will maintain her collection of images on her computer instead of in file cabinets. This has the advantage that it is easier to keep off-site backups, easier to access the images remotely from other labs, easier to organize the files, and so on. (The scientist *must* be careful to use a sophisticated backup system to avoid career-destroying disk-drive crashes!)

The image contained in an electronic document may be at very high resolution. But editors and referees may simply view the image on the good-quality color monitors of their computers; or the referee could print the image on a medium-quality color printer. Though this will not provide a very high-resolution view, it may be sufficient for the job of verifying the scientific validity of the paper. Then, when the journal's production department gets the electronic manuscript, they can photoengrave it at very high resolution.

Electronic distribution

Electronic submission and refereeing of TOPLAS articles has worked smoothly and efficiently since 1993. It is tempting to conclude that electronic production and distribution of articles to subscribers will also be easy. Unfortunately, this does not follow. The Association for Computing Machinery—the society that publishes TOPLAS and a dozen other computer science journals—has been working on an electronic publishing plan [5] since 1992; they have made good progress, but the road has been bumpy. It is not obvious how to translate many different document formats used by authors into a common typesetting language. For refereeing, an author can produce a readable (but not editable) Postscript file no matter what word processor she is using, but professional copy editing and typesetting requires compatibility at the level of the word processor—this is more difficult.

Conclusion

The speed of Internet e-mail, its ability to ship entire documents, its interactivity (allowing quick and convenient responses to queries), and the ability to attach tracking keywords to messages makes for a simple and effective automated editorial traffic management system. The system can be made reasonably secure with firewalls and encrypted e-mail, and should be able to accommodate digitally encoded photographs and multimedia.

References

[1] Andrew W. Appel. How to edit a journal by e-mail. *Journal of Scholarly Publishing*, 27(2):82–99, January 1996.

[2] Andre Bacard. Anonymous remailer FAQ. http://www.well.com/user/abacard/remail.html, 1996.

[3] Andre Bacard. *The Computer Privacy Handbook*. Peachpit Press, Berkeley, California, 1996.

[4] William Cheswick and Steven Bellovin. *Firewalls and Internet Security: Repelling the Wily Hacker*. Addison Wesley, Reading, MA, 1994.

[5] Peter J. Denning and Bernard Rous. The ACM electronic publishing plan. *Communications of the ACM*, 38(4):97–103, 1995.

[6] Simson Garfinkel. *PGP: Pretty Good Privacy*. O'Reilly and Associates, Sebastopol, CA, 1994.

[7] C. Partridge. Mail routing and the domain system. Technical Report Internet Request For Comment 974, Network Information Center, SRI International, Menlo Park, California, February 1986.