Peter J. Denning

Editorial:
Plagiarism in the Web

About a year ago the referees of the EURO PAR 95 conference reported shocking similarities between the work submitted by a particular author and work published previously by well-known researchers. They undertook an investigation in which they concluded the author had committed at least 13 acts of plagiarism over the past several years. By checking with the original authors, they confirmed the plagiarist (nicknamed hereafter as "Mr. X") had already *published* seven articles that had previously been published under different titles and authorship, and had *submitted* six more of similar appropriation. They said Mr. X had apparently retrieved copies of the original technical reports by Internet FTP. Because this was an international incident involving people from several countries, they issued a public report detailing the evidence and calling on all program committees to be on the lookout for Mr. X and all authors who might be victims to take whatever individual actions they deemed appropriate.

What is unusual about this case is not that the referees of a conference detected plagiarism—they often do when it is attempted, which is not very often—but the extent to which the unauthorized copying could be definitively proved. The differ-ence was a tool developed by Narayanan Shivakumar and Hector Garcia-Molina of Stanford University Computer Science Department. This tool, called SCAM (Stanford Copy Analysis Mechanism) was developed as a means of detecting plagiarism, copies, extracts, and strongly similar documents in digital libraries, to support rights-holders in protecting their work against unauthorized copying. SCAM was able to detect at least 13 cases of possible plagiarism involving Mr. X, which were subsequently verified by the original authors.

When the EURO PAR 95 committee report came to ACM, the Publications Board determined three of the articles in question were published under ACM copyright. This meant we could prove Mr. X's published and submitted works were copied without authorization from ACM copyrighted works. That was enough for the Publications Board to act. The Board wished to protect ACM's interest as the copyright holder and to protect the interests of the authors whose works had been copied without authorization and were being wrongfully attributed to another person.

On September 21, 1995, the ACM, through its legal counsel, sent a letter to Mr. X informing him we are aware he had copied three articles without permission and had claimed to be author or coauthor. ACM demanded Mr. X issue an apology, destroy all ACM copyright materials in his possession, and agree to not repeat this in the future. Mr. X has sent ACM a letter indicating compliance with these demands. ACM has notified the original authors of the articles Mr. X copied.

If there is a moral to this sad story, it is that a would-be plagiarist can no longer rely on the immense size of the computing literature to avoid detection. The literature is being stored in digital libraries, where it can be scanned automatically by copy detectors (such as SCAM). It will not be long before we make it a standard part of the editorial processing of a submitted article to include a copy detection test against the computing literature stored in digital libraries.

ACM considers unauthorized copying to be a serious breach and will pursue anyone attempting it with ACM literature. Authors who transfer copyright to ACM can rest assured ACM will do everything in its power to track down such offenders. We will vigorously protect the integrity of our scientific literature, and we will protect our authors' names and works from unauthorized attribution to others. C

Peter Denning is chair of the ACM Publications Board.