Preface

The PAN workshop brought together experts and researchers around the exciting and future-oriented topics of plagiarism detection, authorship identification, and the detection of social software misuse. The development of new solutions for these problems can benefit from the combination of existing technologies, and in this sense the workshop provides a platform that spans different views and approaches.

Plagiarism analysis is a collective term for computer-based methods to identify a plagiarism offense. In connection with text documents we distinguish between corpus-based and intrinsic analysis: the former compares suspicious documents against a set of potential original documents, the latter identifies potentially plagiarized passages by analyzing the suspicious document with respect to changes in writing style.

Authorship identification divides into so-called attribution and verification problems. In the authorship attribution problem, one is given examples of the writing of a number of authors and is asked to determine which of them authored given anonymous texts. In the authorship verification problem, one is given examples of the writing of a single author and is asked to determine if given texts were or were not written by this author. As a categorization problem, verification is significantly more difficult than attribution. Authorship verification and intrinsic plagiarism analysis represent two sides of the same coin.

"Social Software Misuse" can nowadays be noticed on many social software based platforms. These platforms like Blogs, sharing sites for photos and videos, wikis and on-line forums are contributing up to one third of new Web content. "Social Software Misuse" is a collective term for anti-social behavior in on-line communities; an example is the distribution of spam via the e-mail infrastructure. Interestingly, spam is one of the few misuses for which detection technology is developed at all, though various forms of misuse exist that threaten the different on-line communities. Our workshop shall close this gap and invites contributions concerned with all kinds of social software misuse.

September 2009

Benno Stein
Paolo Rosso
Efstathios Stamatatos
Moshe Koppel
Eneko Agirre