Masterarbeit

Karlsruher Institut für Technologie

Institut für Programmstrukturen und Datenorganisation (IPD)
Lehrstuhl für Systeme der Informationsverwaltung, Prof. Böhm

Masterarbeit

Automatic De-anonymization of Scientific Publications

In science, the peer-reviewed publication process often is "double-blind", i.e., the identity and affiliation of authors is hidden from the reviewer, and so are the reviewers from the authors. There are several good intentions behind this: avoiding discrimination, tempering conflicts of interest, or helping reviewers to write impartial reviews.

However, the guarantees brought by the double-blind system often are illusory. In many cases, the reviewers can easily find out who the authors are: A paper is often linked to previous work of its authors or uses specific software and data identifying them. Also, some of its content may appear in previously published preprints or technical reports. Reviewers are often asked not to try to de-anonymize the paper, but it is virtually impossible to enforce such behaviour.

Protected by their anonymity, non-zealous or busy reviewers may then be influenced by partial information which they are assumed not to have. Then the outcome of the review process tends to be unfair. In this work, we want to design algorithms to de-anonymize papers with very high accuracy. The goal is to point out to what extent the concept of anonymity in double-blind reviewing holds. On the other hand, a paper that is not easy to de-anonymize might indicate novelty.

There are previous attempt to address this task [1, 2]. However, the existing work typically base on custom feature extraction or overly complex methods, while their results show a large margin for improvement. One idea here is to improve those results by proposing a new semantic analysis based on recent text embedding models [3] (e.g., as in [4]) together with citation and phrase analysis, to establish a new state-of-the-art in paper de-anonymization.

This results in the following tasks:

- Literature review focusing on de-anonymization methods.
- Proposal and implementation of a new approach to de-anonymize virtually any scientific paper.
- Data collection from scientific papers published, e.g., from here[a] or here [b].
- Extensive experiments to show the success of the method and superiority w.r.t. existing work[c].

To successfully conduct this thesis project, the student must possess:

- Knowledge of Python or Scala programming. A strong interest in Data Mining.
- The ability to plan and work independently. A working knowledge of English.
- A high level of motivation, enthusiasm and curiosity.

To help you with this task, we offer:

- Thorough mentoring and recurrent meetings with your advisor.
- Access to our institute's computing infrastructure (if required).
- Help to "go for the extra mile" and publish parts of your results in scientific conferences.

Throughout this work, the student will acquire knowledge and practical experience in Data Mining, and make a real contribution to the discussion on how peer-reviewing must be carried out.

[1] M. Payer et al. "What You Submit Is Who You Are: A Multimodal Approach for Deanonymizing Scientific Publications". In: *IEEE Trans. Inf. Forensics Secur.* 10.1 (2015), pp. 200–212.
[2] C. Zhang et al. "Camel: Content-Aware and Meta-path Augmented Metric Learning for Author Identification". In: *WWW*. ACM, 2018, pp. 709–718.
[3] Y. Meng et al. "Spherical Text Embedding". In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada*. Ed. by H. M. Wallach et al. 2019, pp. 8206–8215.
[4] E. Fouché et al. "Mining Text Outliers in Document Directories". In: *ICDM*. In press. IEEE Computer Society, 2020.

[a] `https://www.aminer.org/citation`
[b] `https://www.kaggle.com/Cornell-University/arxiv`
[c] `https://github.com/chuxuzhang`

Ansprechpartner

Dr.-Ing. Edouard Fouché | edouard.fouche@kit.edu | Raum: 342
Am Fasanengarten 5 | 76131 Karlsruhe | Gebäude: 50.34