Venue Recommendation: Submitting your Paper with Style
Zaihan Yang and Brian D. Davison
Department of Computer Science and Engineering, Lehigh University

Motivation & Problem Definition
- The quantity and variety of publications have grown in recent decades such that we now have publications across many different topics, genres and writing formats.
- SIGIR vs. SIGMOD; J.ASIST (journal) vs. JCDL (conference); ICML vs. ICMIA
- Researchers have the problem of determining where to submit their finished paper.
- Is there an automatic mechanism in helping to predict or provide recommendations to researchers on their paper submissions?
- Problem Definition:
  - Given a paper, with its information (title, abstract, full content, and references) provided, predict the real publishing venue of this paper, or recommend a list of possible venues that this paper can consider to submit.
  - A Ranking Problem: $P(v | p)$

Methodology
- Basic Collaborative-filtering (CF) [1,2] for venue recommendation
  - Paper Similarity
  - Cosine similarity
  - Content feature vector:
    - <Topic distribution over 100 topics> LDA

Conclusions
- We proposed an effective collaborative-filtering based approach, as demonstrated by experiment results, to predict the real venue publication of a given paper.
- Incorporating stylometric features can improve prediction results.
- Differentiating the importance of different categories of neighboring nodes can further improve the performance.

Contributions
- We proposed a modified collaborative-filtering based approach in fulfilling this task.
- Two extensions:
  - Incorporating stylometric features;
  - Differentiating the importance of different kinds of neighboring nodes;
- We carried out experiments based on real-world historical data that demonstrate the effectiveness of our proposed method.

Experiments
- Data set:
  - ACM: 170,959 papers, 172,890 venues, 2017
  - CiteSeer: 478,805 papers, 110,627 venues, 47,797

- Evaluation:
  - Randomly choose 10000 papers from ACM and CiteSeer dataset.
  - Compare the predicted results with the ground truth.
- Metrics:
  - Accuracy@5,10,20; MRR

Impact of stylometric features
- Incorporating stylometric features can improve performance.

Comparison with other approaches
- Increased by 13.19% (ACM) and 14.01% (CiteSeer) in terms of Accuracy@5.

References