A Study of Academic Collaboration in Computational Linguistics with Latent Mixtures of Authors

Nikhil Johri, Daniel Ramage
Department of Computer Science
Stanford University
Stanford, CA, USA

Daniel A. McFarland
School of Education
Stanford University
Stanford, CA, USA

Daniel Jurafsky
Department of Linguistics
Stanford University
Stanford, CA, USA

{njohri2,dramage,dmcfarla,jurafsky}@stanford.edu

Abstract

Academic collaboration has often been at the forefront of scientific progress, whether amongst prominent established researchers, or between students and advisors. We suggest a theory of the different types of academic collaboration, and use topic models to computationally identify these in Computational Linguistics literature. A set of *author-specific* topics are learnt over the ACL corpus, which ranges from 1965 to 2009. The models are trained on a per year basis, whereby only papers published up until a given year are used to learn that year's author topics. To determine the collaborative properties of papers, we use, as a metric, a function of the cosine similarity score between a paper's term vector and each author's topic signature in the year preceding the paper's publication. We apply this metric to examine questions on the nature of collaborations in Computational Linguistics research, finding that significant variations exist in the way people collaborate within different subfields.

1 Introduction

Academic collaboration is on the rise as single authored work becomes less common across the sciences (Rawlings and McFarland, 2011; Jones et al., 2008; Newman, 2001). In part, this rise can be attributed to the increasing specialization of individual academics and the broadening in scope of the problems they tackle. But there are other advantages to collaboration, as well: they can speed up production, diffuse knowledge across authors, help train new scientists, and are thought to encourage greater innovation. Moreover, they can integrate scholarly communities and foster knowledge transfer between related fields. But all collaborations aren't the same: different collaborators contribute different material, assume different roles, and experience the collaboration in different ways. In this paper, we present a new frame for thinking about the variation in collaboration types and develop a computational metric to characterize the distinct contributions and roles of each collaborator within the scholarly material they produce.

The topic of understanding collaborations has attracted much interest in the social sciences over the years. Recently, it has gained traction in computer science, too, in the form of social network analysis. Much work focuses on studying networks formed via citations (Radev et al., 2009; White and Mccain, 1998), as well as co-authorship links (Nascimento et al., 2003; Liu et al., 2005). However, these works focus largely on the graphical structure derived from paper citations and author co-occurrences, and less on the textual content of the papers themselves. In this work, we examine the nature of academic collaboration using text as a primary component.

We propose a theoretical framework for determining the types of collaboration present in a document, based on factors such as the number of established authors, the presence of unestablished authors and the similarity of the established authors' past work to the document's term vector. These collaboration types attempt to describe the nature of co-authorships between students and advisors (e.g. "apprentice" versus "new blood") as well as those solely between established authors in the field. We present a decision diagram for classifying papers into these types, as well as a description of the intuition behind each collaboration class.

We explore our theory with a computational method to categorize collaborative works into their collaboration types using an approach based on topic modeling, where we model every paper as a latent mixture of its authors. For our system, we use Labeled-LDA (LLDA (Ramage et al., 2009)) to train models over the ACL corpus for every year of the words best attributed to each author in all the papers they write. We use the resulting author signatures as a basis for several metrics that can classify each document by its collaboration type.

We qualitatively analyze our results by examining the categorization of several high impact papers. With consultation from prominent researchers and textbook writers in the field, we demonstrate that our system is able to differentiate between the various types of collaborations in our suggested taxonomy, based only on words used, at low but statistically significant accuracy. We use this same similarity score to analyze the ACL community by sub-field, finding significant deviations.

2 Related Work

In recent years, popular topic models such as Latent Dirichlet Allocation (Blei et al., 2003) have been increasingly used to study the history of science by observing the changing trends in term based topics (Hall et al., 2008), (Gerrish and Blei, 2010). In the case of Hall et al., regular LDA topic models were trained over the ACL anthology on a per year basis, and the changing trends in topics were studied from year to year. Gerrish and Blei's work computed a measure of influence by using Dynamic Topic Models (Blei and Lafferty, 2006) and studying the change of statistics of the language used in a corpus.

These models propose interesting ideas for utilizing topic modeling to understand aspects of scientific history. However, our primary interest, in this paper, is the study of academic collaboration between different authors; we therefore look to learn models for authors instead of only documents. Popular topic models for authors include the Author-Topic Model (Rosen-Zvi et al., 2004), a simple extension of regular LDA that adds an additional author variable over the topics. The Author-Topic Model learns a distribution over words for each topic, as in regular LDA, as well as a distribution over topics for each author. Alternatively, Labeled LDA (Ramage et al., 2009), another LDA variation, offers us the ability to directly model authors as topics by considering them to be the topic labels for the documents they author.

In this work, we use Labeled LDA to directly model probabilistic term 'signatures' for authors. As in (Hall et al., 2008) and (Gerrish and Blei, 2010), we learn a new topic model for each year in the corpus, allowing us to account for changing author interests over time.

3 Computational Methodology

The experiments and results discussed in this paper are based on a variation of the LDA topic model run over data from the ACL corpus.

3.1 Dataset

We use the ACL anthology from years 1965 to 2009, training over 12,908 papers authored by over 11,355 unique authors. We train our per year topic models over the entire dataset; however, when evaluating our results, we are only concerned with papers that were authored by multiple individuals as the other papers are not collaborations.

3.2 Latent Mixture of Authors

Every abstract in our dataset reflects the work, to some greater or lesser degree, of all the authors of that work. We model these degrees explicitly using a latent mixture of authors model, which takes its inspiration from the learning machinery of LDA (Blei et al., 2003) and its supervised variant Labeled LDA (Ramage et al., 2009). These models assume that documents are as a mixture of 'topics,' which themselves are probability distributions over the words in the vocabulary of the corpus. LDA is completely unsupervised, assuming that a latent topic layer exists and that each word is generated from one underlying topic from this set of latent topics. For our purposes, we use a variation of LDA in which we assume each document to be a latent mixture of its *authors*. Unlike LDA, where each document draws a multinomial over all topics, the latent mixture of authors model we use restricts a document to only sample from topics corresponding to its authors. Also, unlike models such as the Author-Topic Model (Rosen-Zvi et al., 2004), where authors are modeled as distributions over latent topics, our model associates each author to exactly one topic, modeling authors directly as distributions over words.

Like other topic models, we will assume a generative process for our collection of D documents from a vocabulary of size V. We assume that each document d has N_d terms and M_d authors from a set of authors A. Each author is described by a multinomial distribution β_a over words V, which is initially unobserved. We will recover for each document a hidden multinomial $\theta^{(d)}$ of length M_d that describes which mixture of authors' best describes the document. This multinomial is in turn drawn from a symmetric Dirichlet distribution with parameter α restrict to the set of authors $\lambda^{(d)}$ for that paper. Each document's words are generated by first picking an author z_i from $\theta^{(d)}$ and then drawing a word from the corresponding author's word distribution. Formally, the generative process is as follows:

- For each author a, generate a distribution β_a over the vocabulary from a Dirichlet prior μ
- For each document d, generate a multinomial mixture distribution $\theta^{(d)} \sim Dir(\alpha.\mathbf{1}_{\lambda^{(d)}})$
- For each document d,
  - For each $i \in \{1, ..., N_d\}$
    - Generate $z_i \in \{\lambda_1^{(d)}, ..., \lambda_{M_d}^{(d)}\} \sim Mult(\theta^{(d)})$
    - Generate $w_i \in \{1, ..., V\} \sim Mult(\beta_{z_i})$

We use Gibbs sampling to perform inference in this model. If we consider our authors as a label space, this model is equivalent to that of Labeled LDA (Ramage et al., 2009), which we use for inference in our model, using the variational objective in the open source implementation[1]. After inference, our model discovers the distribution over terms that best describes that author's work in the presence of other authors. This distribution serves as a 'signature' for an author and is dominated by the terms that author uses frequently across collaborations. It is worth noting that this model constrains the learned 'topics' to authors, ensuring directly interpretable results that do not require the interpretation of a latent topic space, such as in (Rosen-Zvi et al., 2004).

To imbue our model with a notion of time, we train a separate LLDA model for each year in the corpus, training on only those papers written before and during the given year. Thus, we have separate 'signatures' for each author for each year, and each signature only contains information for the specific author's work up to and including the given year. Table 1 contains examples of such term signatures computed for two authors in different years. The top terms and their fractional counts are displayed.

4 Studying Collaborations

There are several ways one can envision to differentiate between types of academic collaborations. We focus on three factors when creating collaboration labels, namely:

- Presence of unestablished authors
- Similarity to established authors
- Number of established authors

If an author whom we know little about is present on a collaborative paper, we consider him or her to be a new author. We threshold new authors by the number of papers they have written up to the publication year of the paper we are observing. Depending on whether this number is below or above a threshold value, we consider an author to be *established* or *unestablished* in the given year.

Similarity scores are measured using the trained LLDA models described in Section 3.2. For any given paper, we measure the similarity of the paper to one of its (established) authors by calculating the cosine similarity of the author's signature in the year preceding the paper's publication to the paper's term-vector.

Using the aforementioned three factors, we define the following types of collaborations:

- **Apprenticeship Papers** are authored by one or more established authors and one or more unestablished authors, such that the similarity of the paper to more than half of the established authors is high. In this case, we say that the new author (or authors) was an apprentice of

[1] http://nlp.stanford.edu/software/tmt/

Philipp Koehn, 2002		Philipp Koehn, 2009		Fernando Pereira, 1985		Fernando Pereira, 2009	
Terms	Counts	Terms	Counts	Terms	Counts	Terms	Counts
word	3.00	translation	69.78	grammar	14.99	type	40.00
lexicon	2.00	machine	34.67	phrase	10.00	phrase	30.89
noun	2.00	phrase	26.85	structure	7.00	free	23.14
similar	2.00	english	23.86	types	6.00	grammar	23.10
translation	1.29	statistical	19.51	formalisms	5.97	constraint	23.00
purely	0.90	systems	18.32	sharing	5.00	logical	22.41
accuracy	0.90	word	16.38	unification	4.97	rules	21.72

Table 1: Example term 'signatures' computed by running a Labeled LDA model over authors in the ACL corpus on a per year basis: top terms for two authors in different years are shown alongside their fractional counts.

the established authors, continuing in their line of work.

- **New Blood Papers** are authored by one established author and one or more unestablished authors, such that the similarity of the paper to the established author is low. In this case, we say that the new author (or authors) provided new ideas or worked in an area that was dissimilar to that which the established author was working in.
- **Synergistic Papers** are authored only by established authors such that it does not heavily resemble any authors' previous work. In this case, we consider the paper to be a product of synergy of its authors.
- **Catalyst Papers** are similar to synergistic ones, with the exception that unestablished authors are also present on a Catalyst Paper. In this case, we hypothesize that the unestablished authors were the catalysts responsible for getting the established authors to work on a topic dissimilar to their previous work.

The decision diagram in Figure 1 presents an easy way to determine the collaboration type assigned to a paper.

5 Quantifying Collaborations

Following the decision diagram presented in Figure 1 and using similarity scores based on the values returned by our latent author mixture models (Section 3.2), we can deduce the collaboration type to assign to any given paper. However, absolute categorization requires an additional thresholding of author similarity scores. To avoid the addition of an arbitrary threshold, instead of directly categorizing papers, we rank them based on the calculated similarity scores on three different spectra. To facilitate ease of interpretation, the qualitative examples we present are drawn from high PageRank papers as calculated in (Radev et al., 2009).

5.1 The MaxSim Score

To measure the similarity of authors' previous work to a paper, we look at the cosine similarity between the term vector of the paper and each author's term signature. We are only interested in the highest cosine similarity score produced by an author, as our categories do not differentiate between papers that are similar to one author and papers that are similar to multiple authors, as long as high similarity to any single author is present. Thus, we choose our measure, the MaxSim score, to be defined as: $\max_{a \in est} cos(a_{sig}, paper)$

We choose to observe the similarity scores only for established authors as newer authors will not have enough previous work to produce a stable term signature, and we vary the experience threshold by year to account for the fact that there has been a large increase in the absolute number of papers published in recent years.

Depending on the presence of new authors and the number of established authors present, each paper can be placed into one of the three spectra: the Apprenticeship-New Blood spectrum, the Synergy spectrum and the Apprenticeship-Catalyst spectrum. Apprenticeship and Low Synergy papers are those with high MaxSim scores, while low scores indicate New Blood, Catalyst or High Synergy papers.

5.2 Examples

The following are examples of high impact papers as they were categorized by our system:

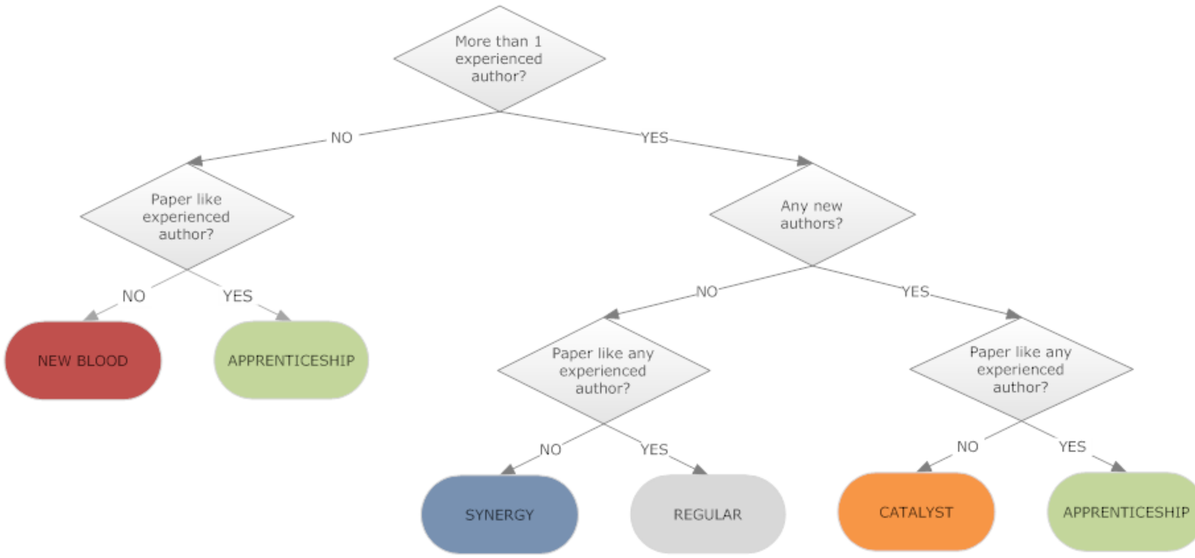

Figure 1: Decision diagram for determining the collaboration type of a paper. A minimum of 1 established author is assumed.

5.2.1 Example: Apprenticeship Paper

Improvements in Phrase-Based Statistical Machine Translation (2004)
by Richard Zens and Hermann Ney
This paper had a high MaxSim score, indicating high similarity to established author Hermann Ney. This categorizes the paper as an Apprenticeship Paper.

5.2.2 Example: New Blood Paper

Thumbs up? Sentiment Classification using Machine Learning Techniques (2002)
by Lillian Lee, Bo Pang and Shivakumar Vaithyanathan
This paper had a low MaxSim score, indicating low similarity to established author Lillian Lee. This categorizes the paper as a New Blood Paper, with new authors Bo Pang and Shivakumar Vaithyanathan. It is important to note here that new authors do not necessarily mean young authors or grad students; in this case, the third author on the paper was experienced, but in a field outside of ACL.

5.2.3 Example: High Synergy Paper

Catching the Drift: Probabilistic Content Models, with Applications to Generation and Summarization (2003)
by Regina Barzilay and Lillian Lee
This paper had low similarity to both established authors on it, making it a highly synergistic paper. Synergy here indicates that the work done on this paper was mostly unlike work previously done by either of the authors.

5.2.4 Example: Catalyst Paper

Answer Extraction (2000)
by Steven Abney, Michael Collins, Amit Singhal
This paper had a very low MaxSim score, as well as the presence of an unestablished author, making it a Catalyst Paper. The established authors (from an ACL perspective) were Abney and Collins, while Singhal was from outside the area and did not have many ACL publications. The work done in this paper focused on information extraction, and was unlike that previously done by either of the ACL established authors. Thus, we say that in this case, Singhal played the role of the catalyst, getting the other two authors to work on an area that was outside of their usual range.

5.3 Evaluation

5.3.1 Expert Annotation

To quantitatively evaluate the performance of our system, we prepared a subset of 120 papers from among the highest scoring collaborative papers based on the PageRank metric (Radev et al., 2009). Only those papers were selected which had at least a single established author. One expert in the field was asked to annotate each of these papers as being either similar or dissimilar to the established authors' prior work given the year of publication, the title of the publication and its abstract.

We found that the MaxSim scores of papers labeled as being similar to the established authors were, on average, higher than those labeled as dissimilar. The average MaxSim score of papers annotated as low MaxSim collaboration types (High Synergy, New Blood or Catalyst papers) was 0.15488, while that of papers labeled as high MaxSim types (Apprentice or Low Synergy papers) had a mean MaxSim score of 0.21312. The MaxSim scores of the different sets were compared using a t-test, and the difference was found to be statistically significant with a two-tailed p-value of 0.0041.

Framing the task as a binary classification problem, however, did not produce very strong results. The breakdown of the papers and success rates (as determined by a tuned threshold) can be seen in Table 3. The system had a relatively low success rate of 62.5% in its binary categorization of collaborations.

5.3.2 First Author Prediction

Studies have suggested that authorship order, when not alphabetical, can often be quantified and predicted by those who do the work (Sekercioglu, 2008). Through a survey of all authors on a sample of papers, Slone (1996) found that in almost all major papers, "the first two authors are said to account for the preponderance of work". We attempt to evaluate our similarity scores by checking if they are predictive of first author.

Though similarity to previous work is only a small contributor to determining author order, we find that using the metric of cosine similarity between author signatures and papers performs significantly better at determining the first author of a paper than random chance. Of course, this feature alone isn't extremely predictive, given that it's guaranteed to give an incorrect solution in cases where the first author of a paper has never been seen before. To solve the problem of first author prediction, we would have to combine this with other features. We chose two other features - an alphabetical predictor, and a predictor based on the frequency of an author appearing as first author. Although we don't show the regression, we do explore these two other features and find that they are also predictive of author order.

Predictor Feature	**Accuracy**
Random Chance	37.35%
Author Signature Similarity	45.23%
Frequency Estimator	56.09%
Alphabetical Ordering	43.64%

Table 2: Accuracy of individual features at predicting the first author of 8843 papers

Table 2 shows the performance of our prediction feature alongside the others. The fact that it beats random chance shows us that there is some information about authorial efforts in the scores we have computed.

6 Applications

A number of questions about the nature of collaborations may be answered using our system. We describe approaches to some of these in this section.

6.1 The Hedgehog-Fox Problem

From the days of the ancient Greek poet Archilochus, the Hedgehog-Fox analogy has been frequently used (Berlin, 1953) to describe two different types of people. Archilochus stated that "The fox knows many things; the hedgehog one big thing." A person is thus considered a 'hedgehog' if he has expertise in one specific area and focuses all his time and resources on it. On the other hand, a 'fox' is a one who has knowledge of several different fields, and dabbles in all of them instead of focusing heavily on one.

We show how, using our computed similarity scores, one can discover the hedgehogs and foxes of Computational Linguistics. We look at the top 100 published authors in our corpus, and for each author, we compute the average similarity score the author's signature has to each of his or her papers. Note that we start taking similarity scores into account only after an author has published 5 papers, thereby allowing the author to stablize a signature in the corpus and preventing the signature from being boosted by early papers (where author similarity would be artificially high, since the author was new).

We present the authors with the highest average similarity scores in Table 4. These authors can be

Collaboration Type	True Positives	False Positives	Accuracy
New Blood, Catalyst or High Synergy Papers	43	23	65.15%
Apprentice or Low Synergy Papers	32	22	59.25%
Overall	75	45	62.50%

Table 3: Evaluation based on annotation by one expert

considered the hedgehogs, as they have highly stable signatures that their new papers resemble. On the other hand, Table 5 shows the list of foxes, who have less stable signatures, presumably because they move about in different areas.

Author	Avg. Sim. Score
Koehn, Philipp	0.43456
Pedersen, Ted	0.41146
Och, Franz Josef	0.39671
Ney, Hermann	0.37304
Sumita, Eiichiro	0.36706

Table 4: Hedgehogs - authors with the highest average similarity scores

Author	Avg. Sim. Score
Marcus, Mitchell P.	0.09996
Pustejovsky, James D.	0.10473
Pereira, Fernando C. N.	0.14338
Allen, James F.	0.14461
Hahn, Udo	0.15009

Table 5: Foxes - authors with the lowest average similarity scores

6.2 Similarity to previous work by sub-fields

Based on the different types of collaborations discussed in, a potential question one might ask is which sub-fields are more likely to produce *apprentice* papers, and which will produce *new blood* papers. To answer this question, we first need to determine which papers correspond to which sub-fields. Once again, we use topic models to solve this problem. We first filter out a subset of the 1,200 highest page-rank collaborative papers from the years 1980 to 2007. We use a set of topics built by running a standard LDA topic model over the ACL corpus, in which each topic is hand labeled by experts based on the top terms associated with it. Given these topic-term distributions, we can once again use the cosine similarity metric to discover the highly associated topics for each given paper from our smaller subset, by choosing topics with cosine similarity above a certain threshold δ (in this case 0.1).

Topic	Score
Statistical Machine Translation	0.2695
Prosody	0.2631
Speech Recognition	0.2511
Non-Statistical Machine Translation	0.2471
Word Sense Disambiguation	0.2380

Table 6: Topics with highest MaxSim scores (papers are more similar to the established authors' previous work)

Topic	Score
Question Answering	0.1335
Sentiment Analysis	0.1399
Dialog Systems	0.1417
Spelling Correction	0.1462
Summarization	0.1511

Table 7: Topics with lowest MaxSim scores (papers are less similar to the established authors' previous work)

Once we have created a paper set for each topic, we can measure the 'novelty' for each paper by looking at their MaxSim score. We can now find the average MaxSim score for each topic. This average similarity score gives us a notion of how similar to the established author (or authors) a paper in the sub field usually is. Low scores indicate that new blood and synergy style papers are more common, while higher scores imply more non-synergistic or apprenticeship style papers. This could indicate that topics with lower scores are more open ended, while those with higher scores require more formality or training. The top five topics in each category are shown in Tables 6 and 7. The scores of the papers from the two tables were compared using a t-test, and the difference in the scores of the two tables was found to be very statistically significant with a two-tailed p value $<< 0.01$.

7 Discussion and Future Work

Once we have a robust way to score different kinds of collaborations in ACL, we can begin to use these scores as a quantitative tool to study phonemena in the computational linguistics community. With our current technique, we discovered a number of negative results; however, given that our accuracy in binary classification of categories is relatively low, we cannot state for sure whether these are true negative results or a limitation of our model.

7.1 Tentative Negative Results

Among the questions we looked into, we found the following results:

- There was no signal indicating that authors who started out as new blood authors were any more or less likely to survive than authors who started out as apprentices. Survival was measured both by the number of papers eventually published by the author as well as the year of the author's final publication; however, calculations by neither measure correlated with the MaxSim scores of the authors' early papers.

- Each author in the corpus was labeled for gender. Gender didn't appear to differentiate how people collaborated. In particular, there was no difference between men and women based on how they started their careers. Women and men are equally likely to begin as new blood authors as they are to begin as apprentices.

- On a similar note, established male authors are equally likely to partake in new blood or apprentice collaborations as their female counterparts.

- No noticeable difference existed between average page rank scores of a certain categorization of collaborative papers (e.g. high synergy papers vs. low synergy papers).

It is difficult to conclusively demonstrate negative results, particularly given that our MaxSim scores are by themselves not particularly strong discriminators in the binary classification tasks. We consider these findings to be tentative and an opportunity to explore in the future.

8 Conclusion

Not everything we need to know about academic collaborations can be found in the co-authorship graph. Indeed, as we have argued, not all types of collaborations are equal, as embodied by differing levels of seniority and contribution from each co-author. In this work, we have taken a first step toward computationally modeling these differences using a latent mixture of authors model and applied it to our own field, Computational Linguistics. We used the model to examine how collaborative works differ by authors and subfields in the ACL anthology. Our model quantifies the extent to which some authors are more prone to being 'hedgehogs,' whereby they heavily focus on certain specific areas, whilst others are more diverse with their fields of study and may be analogized with 'foxes.'

We also saw that established authors in certain subfields have more deviation from their previous work than established authors in different subfields. This could imply that the former fields, such as 'Sentiment Analysis' or 'Summarization,' are more open to new blood and synergistic ideas, while other latter fields, like 'Statistical Machine Translation' or 'Speech Recognition' are more formal or require more training. Alternatively, 'Summarization' or 'Sentiment Analysis' could just still be younger fields whose language is still evolving and being influenced by other subareas.

This work takes a first step toward a new way of thinking about the contributions of individual authors based on their network of areas. There are many design parameters that still exist in this space, including alternative text models that take into account richer structure and, hopefully, perform better at discriminating between the types of collaborations we identified. We intend to use the ACL anthology as our test bed for continuing to work on textual models of collaboration types. Ultimately, we hope to apply the lessons we learn on modeling this familiar corpus to the challenge of answering large-scale questions about the nature of collaboration as embodied by large scale publication databases such as ISI and Pubmed.

Acknowledgments

This research was supported by NSF grant NSF-0835614 CDI-Type II: What drives the dynamic creation of science? We thank our anonymous reviewers for their valuable feedback and the members of the Stanford Mimir Project team for their insights and engagement.

References

Isaiah Berlin. 1953. *The hedgehog and the fox: An essay on Tolstoy's view of history*. Simon & Schuster.

David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, pages 113–120, New York, NY, USA. ACM.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022.

Sean M. Gerrish and David M. Blei. 2010. A language-based approach to measuring scholarly impact. In *Proceedings of the 26th International Conference on Machine Learning*.

David Hall, Daniel Jurafsky, and Christopher D. Manning. 2008. Studying the history of ideas using topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '08, pages 363–371, Stroudsburg, PA, USA. Association for Computational Linguistics.

B. F. Jones, S. Wuchty, and B. Uzzi. 2008. Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322:1259–1262, November.

Xiaoming Liu, Johan Bollen, Michael L. Nelson, and Herbert Van de Sompel. 2005. Co-authorship networks in the digital library research community. *Information Processing & Management*, 41(6):1462 – 1480. Special Issue on Infometrics.

Mario A. Nascimento, Jörg Sander, and Jeffrey Pound. 2003. Analysis of sigmod's co-authorship graph. *SIGMOD Rec.*, 32:8–10, September.

M. E. J. Newman. 2001. From the cover: The structure of scientific collaboration networks. *Proceedings of the National Academy of Science*, 98:404–409, January.

Dragomir R. Radev, Pradeep Muthukrishnan, and Vahed Qazvinian. 2009. The acl anthology network corpus. In *Proceedings of the 2009 Workshop on Text and Citation Analysis for Scholarly Digital Libraries*, NLPIR4DL '09, pages 54–61, Stroudsburg, PA, USA. Association for Computational Linguistics.

Daniel Ramage, David Hall, Ramesh Nallapati, and Christopher D. Manning. 2009. Labeled lda: a supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1*, EMNLP '09, pages 248–256.

Craig M. Rawlings and Daniel A. McFarland. 2011. Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. *Social Science Research*, 40(3):1001 – 1017.

Michal Rosen-Zvi, Thomas Griffiths, Mark Steyvers, and Padhraic Smyth. 2004. The author-topic model for authors and documents. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, UAI '04, pages 487–494.

Cagan H. Sekercioglu. 2008. Quantifying coauthor contributions. *Science*, 322(5900):371.

RM Slone. 1996. Coauthors' contributions to major papers published in the ajr: frequency of undeserved coauthorship. *Am. J. Roentgenol.*, 167(3):571–579.

Howard D. White and Katherine W. Mccain. 1998. Visualizing a discipline: An author co-citation analysis of information science. *Journal of the American Society for Information Science*, 49:1972–1995.