Measuring the Significance of Writing Style for Recommending Where to Publish – A Case Study

Oscar Karnalim[1], *Aulia Zahrina Qashri*[2]

[1]*Maranatha Christian University, Surya Sumantri Street No 65, Bandung, 40164, Indonesia*
[2]*Karolinska Institutet, Solnavägen 1, Solna, 17177, Sweden*
E-mails: oscar.karnalim@it.maranatha.edu   aulia.zahrina.qashri@stud.ki.se

Abstract: *Writing style plays a role in publication venue recommendation. However, such finding should be observed further; it is concluded from an arbitrary dataset which contains various topics and writing quality. This paper aims to observe that style's impact in a more controlled environment. A dataset with the same specific topic and writing quality was used and analysed. In our case, the dataset is comprised of papers published on reputable software engineering publication venues with natural language generation as the specific topic. According to our observation, writing style only affects significantly on paper title wherein the impact is proportional to n in n-gram. Moreover, the style's impact becomes more salient when the venues are grouped per publisher or only a specific publication type is considered.*

Keywords: *Recommender system, writing style, publication venues, N-gram language model.*

1. Introduction

In academia, Information Technology (IT) plays a significant role in terms of reducing human workload. It automates a lot of manual tasks, which are typically exhaustive and time-consuming. Some of those tasks are reexplaining course materials (which could be helped with the use of educational tools [1]), assisting students to complete their laboratory tasks [2], grading essay assessments [3], detecting unethical activities (such as source code plagiarism [4]), and conducting research.

IT helps in many ways in terms of doing research. Publication venue recommendation [5] is one of the examples. Instead of manually search for the most suitable publication venue, the prospective author is only required to provide their paper (or author data) and let the system recommends a suitable venue. Other examples are paper recommendation [6], automated keyphrases (or index terms) generation [7], and paper search on digital libraries [8]. Among aforementioned supports, publication venue recommendation is argued to be crucial for academics;

the number of publication venues grows rapidly while a considerable number of them are out of author's knowledge.

In publication venue recommendation, writing style is argued to be effective [9, 10]. However, their findings are resulted from a dataset with various topics and writing quality. Further observation is necessary to validate whether the impact is purely caused by writing style.

This paper evaluates the significance of writing style in publication venue recommender systems. To avoid biased result due to various publication topics, various scopes of publication venues and various levels of writing quality, our evaluation dataset is taken from publication venues with the same scope wherein all papers discuss the same specific topic. In such a manner, when perceived only from paper content, writing style is the only remaining factor that plays a role in recommendation. For our case study, the publication venues are taken from software engineering field with natural language generation as the specific topic. To the best of our knowledge, this is the first attempt to evaluate the impact of writing style without considering other factors (i.e., different topics, venues, and writing quality) at the time of the writing.

Findings resulted from this work publication venue recommender systems. Moreover, those findings can also help academics to get higher acceptance rate while submitting a paper. If writing style is considerably important, prior submitting to a venue, the academics should read the venue's previously published works, and prepare their paper in similar writing style. Otherwise, they could only focus on the venue's scope and enhance their writing quality (in addition to providing a clear gap and significant contribution).

This paper is organized as follows. Section 2 describes the related works about recommender systems. Section 3 introduces the methodology used to evaluate whether writing style affects where to publish. Section 4 reports the findings of the evaluation. Finally, Section 5 presents the conclusion and future works.

2. Related works

Recommender system is a program that attempts to recommend the most suitable items (e.g., videos, products, or publication venues) to users based on their interests [11]. When perceived from filtering mechanism, existing recommender systems can be classified in fourfold: content-based filtering, collaborative filtering, demographic filtering, and hybrid filtering [12]. Content-based filtering [13] recommends items based on user's previous choices. Collaborative filtering [14] allows users to give ratings toward some elements and, when such ratings are comprehensive enough, they can be used to recommend items; a user will be recommended with particular items that have been chosen by other users with similar rating behaviour. Demographic filtering [15] takes recommendation based on the principle that users with similar personal attributes (e.g., sex and age range) will have common preferences. Hybrid filtering [15] is a combination of the aforementioned filtering mechanisms. For instance, a work in [16] combines collaborative filtering with the

content-based one. Another example is a work proposed in [17]; it combines collaborative filtering with demographic filtering.

In the context of academic publication, existing recommender systems are focused on suggesting either published works to read [6] or publication venue for disseminating academics' works [5]. Even though most attention is focused on suggesting published works to read, we would argue that suggesting publication venue is more important; academics are more familiarized with published works (especially those that are relevant to their interest) than publication venues.

Generally speaking, resources used by existing publication venue recommender systems can be classified to two categories: paper and author data. Paper data is obtained from the paper that is planned to be published. Yang and Davison [9, 10], for example, utilised the whole paper content in their recommender systems. It is true that considering the whole content can exploit many characteristics hidden in the paper, and some of them can be beneficial for the recommender system. However, as it can be time consuming, several works considered only some parts of it. Medvet, Bartoli, Piccinin [18], for example, relied only on the paper's abstract. This was also followed by Wang et al. [5].

Author data, on the contrary, is more focused on information related to the authors listed on the paper. It is often implemented with an argument that relying on the paper data can be misleading due to text ambiguity [19]. Alhoori, Furuta [20] relied on researcher profile similarity on academic social network. That kind of social relation was also used and observed by Beierle, Tan, Grunert [21]. Another author data that is worth to mention is co-publication data. Luong et al. [19] and Yu et al. [22], for instances, proposed recommender systems that rely on the authors' co-authorship relation. Chen et al. [23] also relied on author data. However, their work was more focused on author-venue relation.

When compared to each other, both of them have their own benefit. Paper data can be more suitable for early-career researchers as they have limited number of previously published papers. Such data can also benefit researchers who have just been involved in a new research track since their previously published papers are not related to their current research. Other than those two, author data can be preferred to avoid text ambiguity provided on paper data.

Several works [9, 10] utilised writing style as a part of paper data for publication venue recommender systems, and they argued that the style improved the performance of such systems. The finding is obtained by evaluating the performance toward a dataset with various topics and writing quality. Consequently, further observation is required to check whether the improved performance is really derived from the use of writing style (without topic and writing quality variation on board), especially on those systems which do not rely on the whole paper content (as such reliance can lead to time consuming process in real environments).

3. Methodology

This paper measures the impact of writing style in a controlled environment. Instead of utilizing an arbitrary dataset for evaluation, it relies only on papers published in

top venues with the same scope (to mitigate the impact of various scopes of publication venues and various levels of writing quality) where all of them cover the same specific topic (to mitigate the impact of various publication topics). In our case, the venues' scope is software engineering and the specific topic is natural language generation. It is important to note that both venues' scope and topic can be replaced as needed since the aim of this paper is to evaluate whether writing style should be considered in publication venue recommender systems and such a style can be derived from any venues' scopes and specific topics. For example, this methodology can be applied on computer science education venues with educational tool as the specific topic.

In general, our methodology consists of two phases: data collection and analysis. Data collection builds the dataset in twofold. At first, top-ranked publication venues were listed from several credible resources. Such resources can be papers related to academic venues and journals with impact factor. In our case, we focus on software engineering venues. Hence, works proposed in [24, 25] were used as the papers related to academic venues while IEEE and ACM software engineering journals with impact factor were used as the journals with impact factor. All venues listed on those resources were then merged where some venues that are inaccessible via the Internet are removed. This sub-phase results in 21 distinct publication venues which consist of 15 conferences and six journals. The details of those venues can be seen in Table 1.

Table 1. Selected publication venues

No	Venue name	Type
1	ACM International Conference on Software Engineering	Conference
2	ACM International Symposium on Software Testing and Analysis	Conference
3	ACM SIGSOFT Symposium on the Foundations of Software Engineering	Conference
4	ACM Transactions on Software Engineering and Methodology	Journal
5	ACM Working Conference on Mining Software Repositories	Conference
6	ACM/IEEE International Symposium on Empirical Software Engineering	Conference
7	Elsevier Information and Software Technology	Journal
8	Elsevier Journal of Systems and Software	Journal
9	IEEE Annual Computer Software and Applications Conference (COMPSAC)	Conference
10	IEEE European Conference on Software Maintenance and Reengineering (CSMR)	Conference
11	IEEE International Conference on Software Maintenance	Conference
12	IEEE International Symposium on Software Reliability Engineering	Conference
13	IEEE International Working Conference on Source Code Analysis and Manipulation	Conference
14	IEEE Software	Journal
15	IEEE Transactions on Software Engineering	Journal
16	IEEE Working Conference on Reverse Engineering	Conference
17	IEEE/ACM International Conference on Automated Software Engineering & Conference	Conference
18	IEEE/ACM International Conference on Program Comprehension	Conference
19	Springer Empirical Software Engineering	Journal
20	Springer International Conference on Fundamental Approaches to Software Engineering	Conference
21	Wiley Software Testing, Verification, and Reliability (STVR)	Journal

Afterwards, papers published on the listed publication venues were crawled through each venue's respective indexing services. For example, papers published on ACM Transactions on Software Engineering and Methodology were crawled from ACM digital library. To guarantee that all crawled papers share the same specific topic, search queries given to all indexing services are similar. Since our targeted specific topic is natural language generation in software engineering, these queries are Natural Language (NL), Natural Language Generation (NLG). Selected papers should have at least one of those queries in their title or abstract. Considering writing style may undergo changes over a long period of time, we limited our selected papers' publication date; they should fall within January 2007 to June 2017. This sub-phase yields 258 papers for our evaluation dataset.

Data analysis revalidates whether writing style actually affects in publication venue recommendation. For our case study, a content-based recommender system was used as an evaluation medium; it is the simplest system that relies solely on the content of used resources (i.e., paper contents in our case).

Three kinds of data for recommendation are used in this evaluation: paper Title (TTL), Abstract (ABS), and Combination of both (COM). The combination of paper title and abstract is considered to validate whether involving more information could lead to other different findings.

In accordance to the kinds of data for recommendation, the system used for evaluation also has three modes: relying only on the papers' title, abstract, or both. We do not consider the whole paper content based on two reasons. First, both title and abstract are argued to be representative toward given paper; they are, by definition, the summary of the whole paper content. Second, considering the whole paper content can be time consuming and therefore can be less practical in real environments.

The recommendation system is based on Information Retrieval and has two components: indexing and retrieving components (Fig. 1). The former converts collected paper data (which are paper titles and/or abstracts) and venue data to publication venue profiles. Whereas, the latter accepts a query (either paper title, abstract, or both) and then retrieves the publication venues based on their relevancy toward the query.

Indexing component develops the profile of each publication venue based on its respective papers. The profile is built in threefold. Firstly, for each paper, the title and/or the abstract (depends on the selected mode) are tokenised to n-gram terms [26] where only alphanumeric terms are considered. These terms are therefore lowercased to nullify the impact of capitalization. Finally, the terms of all papers from the same venue are merged, and considered as the profile of that venue.

Retrieving component ranks the publication venues based on their similarity toward given query (which can be the paper title, abstract, or both depend on the selected mode). The query is tokenised in a similar manner as other papers at indexing component. The similarity is measured using n-gram language model (with n ranged from one to four) [26] since it can capture writing style as a part of document authorship [27]. We do not rely on other algorithms such as vector space model or

latent semantic indexing since they are initially derived from the bag-of-words concept, which can be less contextual compared to the *n*-gram one.
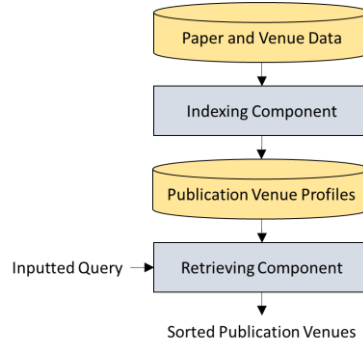

Fig. 1. How indexing and retrieving components in our recommendation system interact to each other. Blue shapes refer to the components; orange shapes refer to the databases, other text refers to either input or output

The impact of writing style was measured through Mean Average Precision (MAP) [26], an effectiveness measurement that is derived from precision. Both of them are calculated from the proportion between correctly predicted results and all predicted results. The only difference is that MAP is more sensitive since it considers rank position on its measurement; relevant results ranked on high position is preferred.

MAP is calculated by considering all papers as the queries and all publication venues as the documents. Each time a paper acts as a query, the paper is excluded from these venues' profiles, and then used to retrieve these venues in descending order. It is important to note that one paper from our dataset cannot be used as a query since its publication venue is exclusive (i.e., it is only assigned to that paper). If that paper acted as a query, the paper's publication venue would have an empty profile as it has no papers.

In addition to MAP, we also evaluate Top-1 MAP (i.e., MAP that is measured only based on the first recommendation) since in most recommender systems, the first recommended result is the most decisive one for user.

To enrich the analysis, three scenarios were observed. The first one is recommending publication venue as it is. It aims to measure the impact of writing style in a default publication venue recommender system. The second one is about publisher recommendation. In this context, we want to measure whether writing style also affects in terms of recommending publishers (i.e., a generalized version of publication venues). The last one is similar to the first except that publication venues are grouped based on their publication type (i.e., conference papers or journal articles). We plan to evaluate whether the writing style favours a particular publication type.

To sum up, three scenarios were conducted. Each scenario was performed under two metrics: *n* in *n*-gram (ranged from one to four) and used data for recommendation (paper title, abstract, or both). Consequently, each of the scenarios led to 12 sub-scenarios.

8

4. Results and discussion

4.1. Writing style on publication venue recommendation

Fig. 2 shows that MAP and Top-1 MAP are proportional to n-gram when paper abstract is used as the data for recommendation. Further observation shows that longer contiguous sequence (provided by applying higher n in n-gram) provides stronger writing style hint in abstract.

MAP and Top-1 MAP on Publication Venue
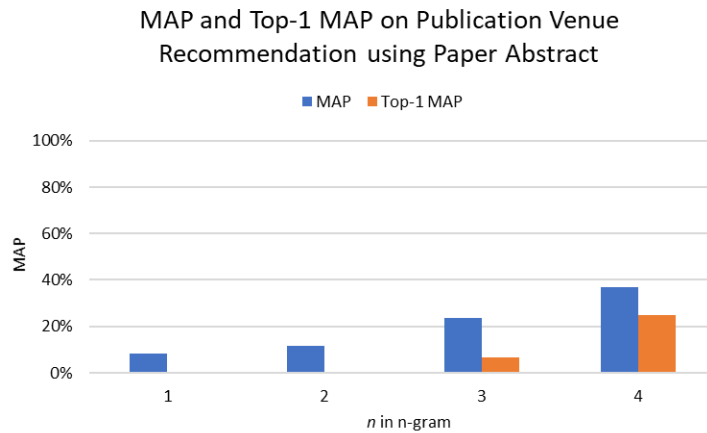Recommendation using Paper Abstract

■ MAP ■ Top-1 MAP

Fig. 2. MAP and Top-1 MAP on publication venue recommendation using paper abstract. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

As seen in Fig. 2, if n is lower or equal to two, Top-1 MAP will be assigned as zero. In other words, writing style is only useful for retrieving the correct venue as the first rank when the sequence contains at least three words.

Generally speaking, writing style on paper abstract is argued to be moderately effective on publication venue recommendation; its highest MAP is only 37.301%. Further, in most occasions, the first recommendation is not the correct one (see Fig. 2 where Top-1 MAP is lower than MAP and the highest Top-1 MAP is still below 25%).

When paper title is used as the data for recommendation (Fig. 3), MAP and Top-1 MAP are still proportional to n in n-gram. Further, the correct venue is not always positioned as the first rank. However, it is more effective than paper abstract for publication venue recommendation. The highest Top-1 MAP is 87%, which is close to the highest possible MAP value. Hence, it can be stated that writing style affects more positively for publication venue recommendation when title is used as the data for recommendation.

The use of both paper title and abstract does not lead to effectiveness improvement. Fig. 4 depicts the effectiveness is considerably similar to publication venue recommendation using paper abstract alone (Fig. 2) with difference less than 1%. Further observation shows that such phenomenon is caused by the existence of paper abstract. The abstract is commonly longer than the title, mitigating the impact of the title in providing recommendation.
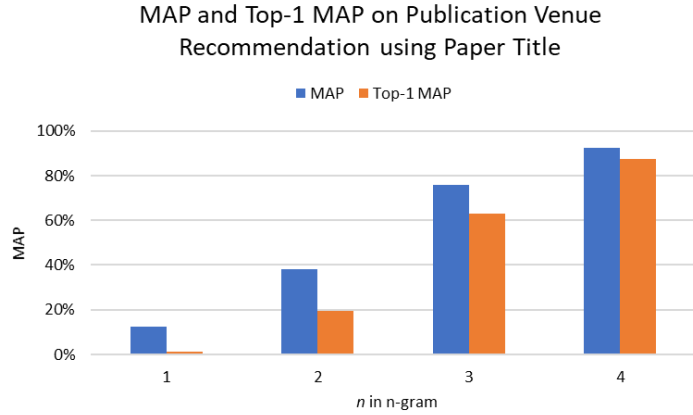
MAP and Top-1 MAP on Publication Venue Recommendation using Paper Title

Fig. 3. MAP and Top-1 MAP on publication venue recommendation using paper title. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

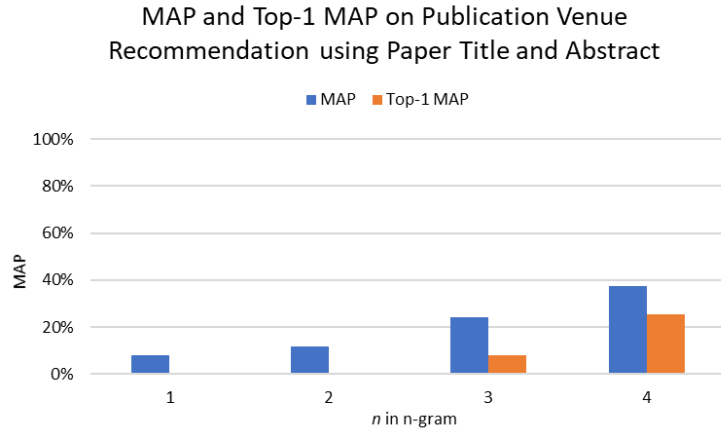**MAP and Top-1 MAP on Publication Venue Recommendation using Paper Title and Abstract**

Fig. 4. MAP and Top-1 MAP on publication venue recommendation using paper title and abstract. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

4.2. Writing style on publisher recommendation

When all publication venues are grouped per publisher, Fig. 5 shows that both MAP and Top-1 MAP are still proportional to n-gram even though MAP leads to higher improvement as n in n-gram is increased. The use of paper abstract is considerably effective (as the highest MAP is 57.59%) even though the correct publisher is only ranked at the first rank on some cases (as Top-1 MAP is still lower than MAP).

When paper title is used as the data for recommendation and n in n-gram is three or four, the system commonly suggests the correct publisher as the first recommendation (see Fig. 6 where the difference between MAP and Top-1 MAP on these sub-scenarios are insignificant). Nevertheless, when n is set as one, the correct publisher is seldom placed as the first recommendation (see the large difference between MAP and Top-1 MAP for n=1 on Fig. 6).

Compared to paper abstract, paper title is more beneficial for publisher recommendation when n in n-gram is higher than one. The title's MAP and Top-1

10

MAP are far higher than the abstract's, with the highest value are 97% for MAP and 96% for Top-1 MAP.
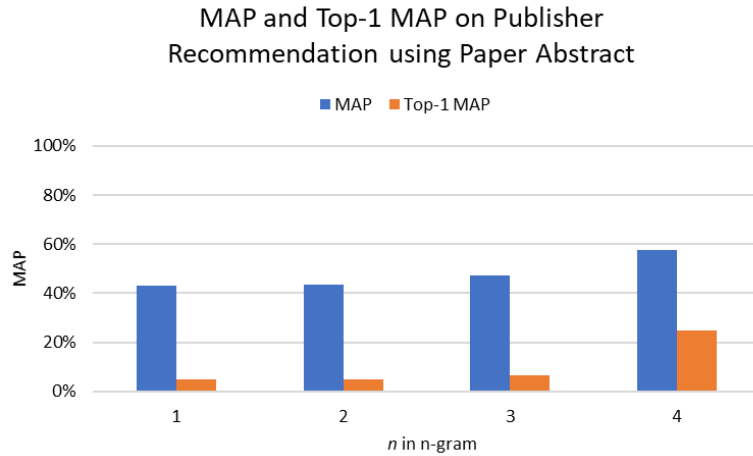
MAP and Top-1 MAP on Publisher Recommendation using Paper Abstract

Fig. 5. MAP and Top-1 MAP on publisher recommendation using paper abstract. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

MAP and Top-1 MAP on Publisher Recommendation using Paper Title
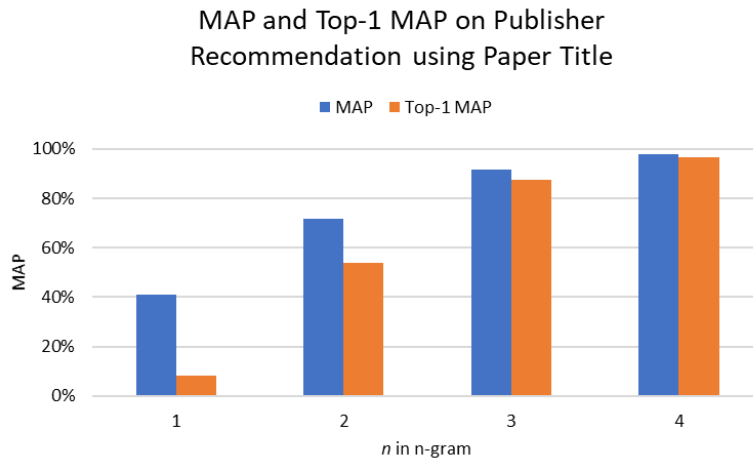

Fig. 6. MAP and Top-1 MAP on publisher recommendation using paper title. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

When combined, the use of paper title and abstract does not lead to higher effectiveness. Fig. 7 depicts that it is as effective as utilizing the paper abstract only. The reason for this is that the title is shorter than the abstract and such imbalance proportion mitigates the impact of the paper title.

Compared to publication venue recommendation, publisher recommendation typically generates higher effectiveness metric. In average, MAP experiences 24% improvement and Top-1 MAP experiences 10.6% improvement. Therefore, writing style has better effect in publisher recommendation.
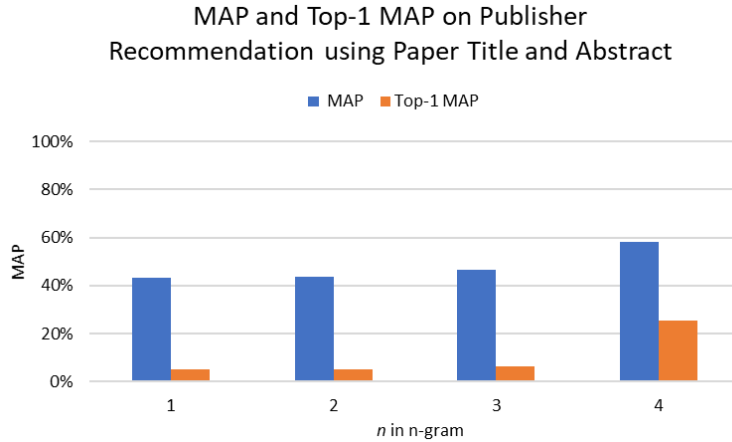
11

MAP and Top-1 MAP on Publisher
Recommendation using Paper Title and Abstract

■ MAP ■ Top-1 MAP

Fig. 7. MAP and Top-1 MAP on publisher recommendation using paper title and abstract. Horizontal axis represents n in n-gram while vertical axis represents resulted MAP

4.3. Writing style on publication venue recommendation grouped by publication type

In this scenario, three approaches are considered: TTL, ABS, and COM. TTL relies on paper title as the data for recommendation while ABS and COM rely on paper abstract and the concatenation of both title & abstract, respectively. According to our previous evaluation scenarios, it is clear that higher n in n-gram generates higher MAP. Hence, to avoid redundant findings, we only consider one n with the largest value (which is four) for all approaches.

Each approach is therefore evaluated toward three data groups: journal articles, conference papers, and mixed papers (i.e., the whole evaluation dataset). Considering the first recommendation is the most crucial one for user and findings related to MAP and Top-1 MAP have been discussed on previous sections, this scenario will only consider Top-1 MAP.

In general, Fig. 8 shows that journal group yields the highest Top-1 MAP, followed by conference and mixed group respectively. Such pattern applies on all types of data for recommendation (i.e., paper title, abstract, or both). Two findings can be stated. First, writing style is more salient on journals than conferences. Unlike conferences, most journals have no tight deadline for reviewing and publishing. The prospective authors will have more time to prepare their paper as the targeted journal's previously published works. Second, grouping based on publication types can boost up effectiveness. Compared to other groups (where each of them only considers one publication type), mixed group yields the lowest Top-1 MAP.

From recommendation resource perspective, the use of paper Title (TTL) still outperforms the use of paper Abstract (ABS) in all groups (See Fig. 8). Writing style provides more impact when the title is used as recommendation resource regardless of publication groups. Another interesting finding from Fig. 8 is that Combining paper title and abstract (COM) shows similar effectiveness as using paper abstract alone (ABS). Further observation shows that the abstract is far longer than the title

12

for all papers and that thing mitigates the impact of the title in publication venue recommendation.

Top-1 MAP on Publication Venue
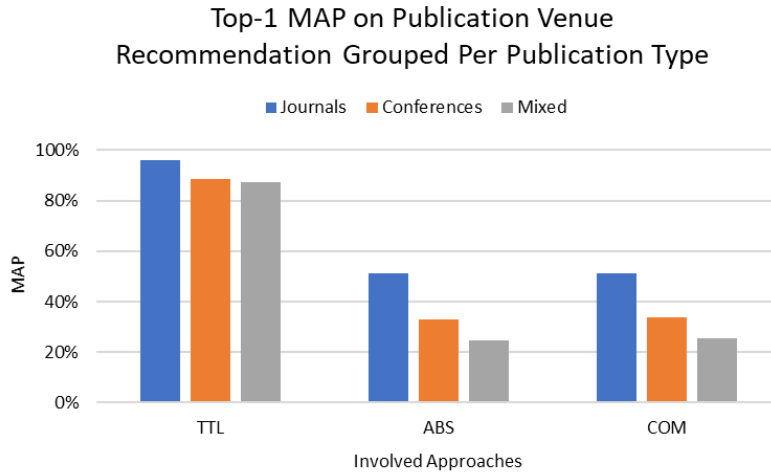Recommendation Grouped Per Publication Type

Fig. 8. Top-1 MAP on publication venue recommendation grouped per publication type. Horizontal axis refers to involved approaches while their results are displayed on vertical axis

5. Conclusion

This paper evaluates the significance of writing style for publication venue recommendation. Four findings can be stated from our evaluation. First, writing style on paper title is crucial for recommendation while similar style on abstract only shows moderate impact. Second, higher n in n-gram accentuates the impact of writing style. Third, writing style's impact is more salient when recommending publication venues' publisher. Fourth, writing style is more effective on journals instead of conferences or the combination of them.

In general, there are two limitations that should be considered toward our findings. First, the dataset is specifically focused on the use of natural language generation in software engineering. Hence, the findings could not be generalised for all available specific topics. It is possible that writing style works in different way when other specific topics are used. Second, resulted findings are derived from n-gram language model. Other recommendation approaches could yield different results.

For future works, we plan to evaluate writing style's impact on other specific topics. Furthermore, we also plan to check whether different recommendation approaches can accentuate the impact of writing style through a comparative study.

R e f e r e n c e s

1. S o r v a, J., V. K a r a v i r t a, L. M a l m i. A Review of Generic Program Visualization Systems for Introductory Programming Education. – ACM Transactions on Computing Education, Vol. **13**, 2013, No 4, p. 15.

2. E l v i n a, E., O. K a r n a l i m, M. A y u b, M. C. W i j a n t o. Combining Program Visualization with Programming Workspace to Assist Students for Completing Programming Laboratory Task. – Journal of Technology and Science Education, Vol. **8**, 2018, No 4, pp. 268-280.

3. M c N a m a r a, D. S., S. A. C r o s s l e y, R. D. R o s c o e, L. K. A l l e n, J. D a i. A Hierarchical Classification Approach to Automated Essay Scoring. – Assessing Writing, Vol. **23**, 2015, No 1, pp. 35-39.

4. K a r n a l i m, O. A Low-Level Structure-Based Approach for Detecting Source Code Plagiarism. – IAENG International Journal of Computer Science, Vol. **44**, 2017, No 4, pp. 501-522.

5. W a n g, D., Y. L i a n g, D. X u, X. F e n g, R. G u a n. A Content-Based Recommender System for Computer Science Publications. – Knowledge-Based Systems, Vol. **157**, 2018, No 1, pp. 1-9.

6. B e e l, J., B. G i p p, S. L a n g e r, C. B r e i t i n g e r. Research-Paper Recommender Systems: A Literature Survey. – International Journal on Digital Libraries, Vol. **17**, 2016, No 4, pp. 305-338.

7. J o n a t h a n, F. C., O. K a r n a l i m. Semi-Supervised Keyphrase Extraction on Scientific Article Using Fact-Based Sentiment. – TELKOMNIKA (Telecommunication Computing Electronics and Control), Vol. **16**, 2018, No 4, pp. 1771-1778.

8. S u n, Y., C. L. G i l e s. Popularity Weighted Ranking for Academic Digital Libraries. – In: Proc. of European Conference on Information Retrieval. Rome, Springer-Verlag, 2007, pp. 605-612.

9. Y a n g, Z., B. D. D a v i s o n. Distinguishing Venues by Writing Styles. – In: Proc. of 12th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM, Washington, 2012, pp. 371-372.

10. Y a n g, Z., B. D. D a v i s o n. Venue Recommendation: Submitting Your Paper with Style. – In: Proc. of 11th International Conference on Machine Learning and Applications, IEEE, Boca Raton, 2012, pp. 681-686.

11. B o b a d i l l a, J., F. O r t e g a, A. H e r n a n d o, A. G u t i é r r e z. Recommender Systems Survey. – Knowledge-Based Systems, Vol. **46**, 2013, No 1, pp. 109-132.

12. A d o m a v i c i u s, G., A. T u z h i l i n. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. – IEEE Transactions on Knowledge and Data Engineering, Vol. **17**, 2005, No 6, pp. 734-749.

13. S a l t e r, J., N. A n t o n o p o u l o s. CinemaScreen Recommender Agent: Combining Collaborative and Content-Based Filtering. – IEEE Intelligent Systems, Vol. **21**, 2006, No 1, pp. 35-41.

14. S u, X., T. M. K h o s h g o f t a a r. A Survey of Collaborative Filtering Techniques. – Advances in Artificial Intelligence, Vol. **2009**, 2009, No 1, pp. 1-19.

15. P o r c e l, C., A. T e j e d a-L o r e n t e, M. A. M a r t í n e z, E. H e r r e r a-V i e d m a. A Hybrid Recommender System for the Selective Dissemination of Research Resources in a Technology Transfer Office. – Information Sciences, Vol. **184**, 2012, No 1, pp. 1-19.

16. B a r r a g á n s-M a r t í n e z, A., E. C o s t a-M o n t e n e g r o, J. C. B u r g u i l l o, M. R e y-L ó p e z, F. A. M i k i c-F o n t e, A. P e l e t e i r o. A Hybrid Content-Based and Item-Based Collaborative Filtering Approach to Recommend TV Programs Enhanced with Singular Value Decomposition. – Information Sciences, Vol. **180**, 2010, No 22, pp. 4290-4311.

17. V o z a l i s, M. G., K. G. M a r g a r i t i s. Using SVD and Demographic Data for the Enhancement of Generalized Collaborative Filtering. – Information Sciences, Vol. **177**, 2007, No 15, pp. 3017-3037.

18. M e d v e t, E., A. B a r t o l i, G. P i c c i n i n. Publication Venue Recommendation Based on Paper Abstract. – In: Proc. of 26th IEEE International Conference on Tools with Artificial Intelligence, IEEE, Limassol, 2014, pp. 1004-1010.

19. L u o n g, H., T. H u y n h, S. G a u c h, L. D o, K. H o a n g. Publication Venue Recommendation Using Author Network's Publication History. – In: Proc. of Asian Conference on Intelligent Information and Database Systems, Springer-Verlag, Kaohsiung, 2012, pp. 426-435.

20. A l h o o r i, H., R. F u r u t a. Recommendation of Scholarly Venues Based on Dynamic User Interests. – Journal of Informetrics, Vol. **11**, 2017, No 2, pp. 553-563.

21. B e i e r l e, F., J. T a n, K. G r u n e r t. Analyzing Social Relations for Recommending Academic Conferences. – In: Proc. of 8th ACM International Workshop on Hot Topics in Planet-Scale Mobile Computing and On-Line Social Networking, ACM, Paderborn, 2016, pp. 37-42.

22. Y u, S., et al. PAVE: Personalized Academic Venue Recommendation Exploiting Co-Publication Networks. – Journal of Network and Computer Applications, Vol. **104**, 2018, No 1, pp. 38-47.

14

23. C h e n, Z., F. X i a, H. J i a n g, H. L i u, J. Z h a n g. AVER: Random Walk Based Academic Venue Recommendation. – In: Proc. of 24th International Conference on World Wide Web, ACM, Florence, 2015, pp. 579-584.

24. C a i, K.-Y., D. C a r d. An Analysis of Research Topics in Software Engineering 2006. – Journal of Systems and Software, Vol. **81**, 2008, No 6, pp. 1051-1058.

25. V a s i l e s c u, B., A. S e r e b r e n i k, T. M e n s. A Historical Dataset of Software Engineering Conferences. – In: Proc. of 10th Working Conference on Mining Software Repositories, IEEE, San Francisco, 2013, pp. 373-376.

26. C r o f t, W. B., D. M e t z l e r, T. S t r o h m a n. Search Engines: Information Retrieval in Practice. Addison-Wesley, 2010.

27. S t a m a t a t o s, E. A Survey of Modern Authorship Attribution Methods. – Journal of the American Society for Information Science and Technology, Vol. **60**, 2009, No 3, pp. 538-556.