ConfAssist: A Conflict resolution framework for assisting the categorization of Computer Science conferences

Mayank Singh, Tanmoy Chakraborty, Animesh Mukherjee, Pawan Goyal
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur, WB, India
{mayank.singh,its_tanmoy,animeshm,pawang}@cse.iitkgp.ernet.in

ABSTRACT

Classifying publication venues into top-tier or non top-tier is quite subjective and can be debatable at times. In this paper, we propose *ConfAssist*, a novel assisting framework for conference categorization that aims to address the limitations in the existing systems and portals for venue classification. We identify various features related to the stability of conferences that might help us separate a top-tier conference from the rest of the lot. While there are many clear cases where expert agreement can be almost immediately achieved as to whether a conference is a top-tier or not, there are equally many cases that can result in a conflict even among the experts. *ConfAssist* tries to serve as an aid in such cases by increasing the confidence of the experts in their decision. A human judgment survey was conducted with 28 domain experts. The results were quite impressive with 91.6% classification accuracy.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Venue classification, Diversity Index, Stability, Conflict resolution

1. INTRODUCTION

The scientific community has always been demanding for better algorithms, metrics and features for scientific venue ranking and categorization. The existing systems and portals for venue classification, however, have several limitations such as no clear demarcation between categories and no description of main intuitions behind such classification. Ranking systems use h-index and impact factor based metrics, which in turn are debatable. Studies on recommending appropriate publication venues to the researcher for their research paper have explored author's network of related coauthors [2, 5] as well as topic and writing-style information [6]. A similar study [4] suggests that the prestige of a venue depends on several factors such as sponsorship by national or international professional organization, reputation of publisher etc. Another study on wellness of software engineering conferences uses features like author and program committee (PC) stability, openness to new authors, scientific prestige etc. [3].

In this paper, we present *ConfAssist* which is a novel conflict resolution framework that can assist experts to resolve conflicts in deciding whether a conference is a top-tier or not by expressing how (dis)similar the conference is to other well accepted top-tier/ non top-tier conferences. This paper tries to answer some of the very pertinent questions: ***1.*** *What are the underlying features behind the popularity of conferences?* ***2.*** *How can these features be meaningfully used to predict the category of a given conference?*

2. DATASET

This paper uses the dataset (see [1]), crawled from Microsoft Academic Search. For our study, we consider papers published from 1999 to 2010 for 92 conferences. A benchmark is built from systems that provide conference categorizations. We consider 4 such systems and compile categories for each of the 92 conferences. A conference is eligible for consideration, if it is present in atleast 3 systems. 73 out of the 92 conferences satisfied this criteria. Out of these 73 conferences, we call a conference as non-conflicting (NC) if it has been labeled using the same category in all the systems, otherwise it is called a conflicting conference (CC). Overall, the set NC contains 37 conferences with 28 labeled as top-tier and 9 labeled as non top-tier.

3. FEATURES AND ANALYSIS

We select 9 different features and study the dynamics of the conferences in terms of how these parameters change over the years.

i) Conference Reference Diversity Index (CRDI) measures how diversified are the fields referred to by the papers published in a conference.

ii) Conference Keyword Diversity Index (CKDI) represents the diversity in the paper keywords.

iii) Conference Author Diversity Index (CADI) corresponds to what fraction of authors with diversified research interests publish in a conference.

iv) Proportion of New Authors (PNA) explores whether the fraction of papers with new authors is roughly the same over the years for the conference.

v) Conference Author Publication Age Diversity Index (CAAI) represents whether the top-tier conferences have more inclination towards maintaining similar publication-age diversity (or diversity in terms of publication experience of the authors) over time.

vi) Degree Diversity Index (DDI) presents the fluctuations in the overall collaborative behavior of the authors in the conference.

vii) Edge Strength Diversity Index (EDI) shows the fluctuations in the choice of the co-authors for a given author in a conference.

viii) Average Closeness centrality (ACC) represents the closeness of an author to other authors in terms of collaboration.

ix) Average Betweenness Centrality (ABC) measures the "importance" of each author in the collaboration network of the authors publishing in a conference.

For each of these 9 features, we use 3 different parameters: the mean, the median and the standard deviation of consecutive year difference (Δ) of raw values. Next, we describe a few experiments conducted to understand the behavior of these features across various categories and fields.

***Comparing top-tier and non top-tier using features*:** Figure 1 presents comparison between INFOCOM (top-tier) and IWQoS (non top-tier) using 3 features, CAAI, DDI and ACC on a yearly scale. One observation is that in majority of features, values are much higher for INFOCOM than for IWQoS. This plot clearly shows that for non top-tier conferences, raw values over the years are very fluctuating, thus they give rise to high mean and standard deviation in the year-wise differences.

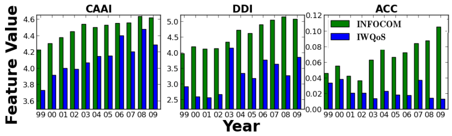

Figure 1: Comparison between INFOCOM (top-tier) and IWQoS (non top-tier) raw feature values using CAAI, DDI and ACC.

***Fieldwise comparison of representative conferences*:** We compare feature values of top-tier conferences with non top-tier in each field. Figure 2 shows plots for 4 computer science fields. For the sake of visualization, we divide our features into three buckets, features 1-9, 10-18 and 19-27. One straightforward observation is that at least for the first two buckets, the feature values for top-tier conferences are lower than those for the non top-tier. On further analysis, we observe that the separation between the two conferences is proportional to the ratio of their respective field ratings. Figure 3 presents plots of two conflicting conferences, ICALP and ECIR, compared against the representative conferences of their fields, *Algorithms and Theory* and *Information Retrieval* respectively. At least for the most discriminative features, the feature values for these conferences lie between the features values of the top-tier and the non top-tier conferences in their field. At the same time, for the medium discriminative features, the values of the conflicting conferences is sometimes even higher than the non top-tier conferences.

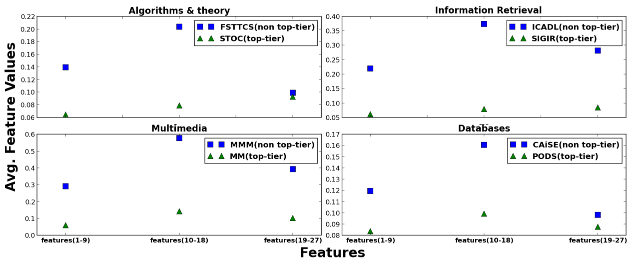

Figure 2: Comparison of feature values of top-tier and non top-tier in four computer science fields.

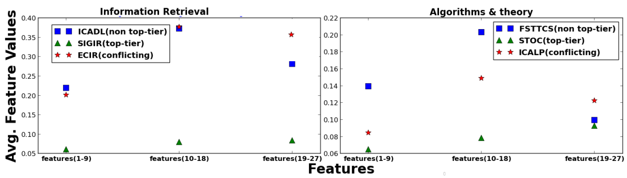

Figure 3: Comparison of feature values of conflicting conferences with top-tier and non top-tier for two CS fields.

***Comparison of newly starting conferences with top-tier and non top-tier*:** We also made an attempt to compare a newly starting conference (for example, JCDL, started in 2001) with top-tier and non top-tier conferences. Figure 4 shows comparison of year-wise profile for JCDL with the average values for all the top-tier and non top-tier using two features, namely ΔPNA and ΔABC. As noted from the left panel, the initial ΔPNA values for JCDL are closely matching with that of the non top-tier conferences. However, after the year 2003, the fluctuation in PNA values is low, even much lower than the average values for the top-tier conferences. The right panel in Figure 4 shows fluctuations in consecutive values of ΔABC. Here, the values for JCDL are below the top-tier average from the beginning. Further analysis on dataset shows that JCDL is slowly promoting the increase in the raw ABC values by allowing higher proportion of bridging authors till 2008, then a sudden drop and again rise in next consecutive years.

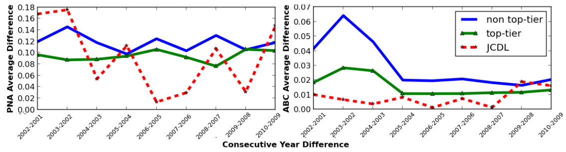

Figure 4: Comparison of year-wise profile for JCDL and average of all top-tier and non top-tier.

4. EXPERIMENTS

The main idea behind this experiment is to predict a category for a conference based on the nearest matching conference from the ground truth dataset. The set NC was taken as the gold-standard to classify any conflicting conference in the set CC. Since the number of top-tier conferences (28) in this set is 3 times the number of non top-tier conferences (9), we created 5 NC sub-sets each having 9 non top-tier and 9 random top-tier conferences. We ran kNN on each of the 5 NC subsets and in each run, we classified a conference in set CC as top-tier, if at-least 4 nearest neighbors are top-tier, otherwise it is classified as a non-top-tier. Finally, a conference is categorized based on the majority from the 5 runs of kNN.

To evaluate the results obtained by our system for this conflicting set, we conduct an online survey[1]. Out of 525 responses, 363 (69.1%) matched our classification results. Considering the majority voting for each conference, 33 out of 36 (91.6%) conferences were correctly classified, 2 were incorrectly classified, while 1 got equal number of matching and non-matching votes. JCDL got 62% votes in favor of top-tier. 148 out of the 226 responses felt more confident about their choice after seeing our results.

5. REFERENCES

[1] T. Chakraborty, S. Kumar, P. Goyal, N. Ganguly, and A. Mukherjee. Towards a stratified learning approach to predict future citation counts. In *Digital Libraries*, 2014.

[2] H. Luong, T. Huynh, S. Gauch, L. Do, and K. Hoang. Publication venue recommendation using author network's publication history. In *Intelligent Information and Database Systems*, pages 426–435. 2012.

[3] B. Vasilescu, A. Serebrenik, T. Mens, M. G. van den Brand, and E. Pek. How healthy are software engineering conferences? *Science of Computer Programming*, 89, Part C:251–272, 2014.

[4] R. E. West and P. J. Rich. Rigor, impact and prestige: A proposed framework for evaluating scholarly publications. *Innovative Higher Education*, 37:359–371, 2012.

[5] F. Xia, N. Y. Asabere, J. J. Rodrigues, F. Basso, N. Deonauth, and W. Wang. Socially-aware venue recommendation for conference participants. In *UIC/ATC*, pages 134–141, 2013.

[6] Z. Yang and B. D. Davison. Venue recommendation: Submitting your paper with style. In *ICMLA*, volume 1, pages 681–686, 2012.

[1]Further Reference and online survey, available at `http://cse.iitkgp.ac.in/resgrp/cnerg/evaluation/JCDL/`.