Incorporating Diversity in Academic Expert Recommendation

Omar Salman, Susan Gauch, Mohammed Alqahatani, Mohammed Ibrahim, Reem Alsaffar
Computer Science and Computer Engineering Department
University of Arkansas
Fayetteville, AR, USA
e-mail:{oasalman, sgauch, ma063, msibrahi, rbalsaff}@uark.edu

***Abstract*—Expert recommendation is the process of identifying individuals who have the appropriate knowledge and skills to achieve a specific task. It has been widely used in the educational environment mainly in the hiring process, paper-reviewer assignment, assembling conference program committees, etc. In this paper, we highlight the problem of diversity and fair representation of underrepresented groups in expertise recommendation, factors that current expertise recommendation systems rarely consider. We present a novel way to model experts in the academic setting by considering the demographic attributes in addition to skills. We use the *h*-index score to quantify skills for a researcher and we identify five demographic features with which to represent a researcher's demographic profile. We highlight the importance of these features and their role in bias within the academic environment. We present three different algorithms for scholar recommendation: expertise-based, diversity-based, and a hybrid approach. To evaluate the ranking produced by these algorithms, we propose a modified normalized Discounted Cumulative Gain (nDCG) version that supports multi-dimensional features and we report the diversity gain from each method. We used a tuning parameter to calibrate the balance between expertise loss and diversity gain. Our results show that we can achieve the best diversity gain increase when the tuning parameter value is set around 0.4, giving nearly equal weight to both expertise and diversity.**

Keywords-Expert Recommendation; Diversity; Fairness; nDCG.

I. INTRODUCTION

We are witnessing a significant change in the amount of the available information. The introduction of social media, blogs, the internet of things, and knowledge sharing communities have dramatically increased the amount of the available knowledge online [1]. This has led modern economies to shift to knowledge-based economies where the intellectual capabilities and expertise of the people determine their values in their enterprise and society [2]. However, determining the level of a person's expertise is a significant challenge because it is quite difficult to assess the amount of knowledge that individuals carry in their minds. Hence, enterprises and companies are beginning to rely on documenting people's expertise, and expert recommendation systems have been developed to identify the right individuals for a task. These systems are mainly dependent on the written artifacts of the experts to determine their expertise. For example, early systems consider the internal documents othe enterprise to extract the skills of individual employees [3].

Expert recommendation systems have been used in academia in the hiring process or finding reviewers or assembling a conference program committee. Although there have been promising developments, most expert recommender systems have not addressed the issue of demographic discrepancies and the need to have a diverse team [3]. Additionally, systems based on machine learning trained on biased training data perpetuate that bias in their recommendations, damaging underrepresented groups [4]. To address this issue, we propose three different approaches with the aim of providing accurate expertise, team diversity, and fair recommendation. Our contribution can be summarized as below:

- Propose a novel way to model an expert in the educational setting using a multivariate profile.
- Present new expert recommendation algorithms that consider different demographic attributes.
- Propose a modified metric that evaluates ranking based on different attributes.

II. LITERATURE REVIW

The process of expertise recommendation has been extensively surveyed by Balog et. al in [3]. The interest in expertise recommendation and expert modeling in academia has been discussed in [3][5]. Although there has been little attention to study an expert demographic profile in academia, Cochran-Smith and Zeichner [6] defined it to include the status of an individual with respect to gender, race, ethnicity, socioeconomic background, and age. Any attempt to model an individual's demographics is complicated due to the fact that people tend not to explicitly provide such information. Hence, there are several approaches to predicting their demographic information from publicly available information such as their name [7]-[13].

Bias in expertise recommendation within academia has received a fair amount of attention by many researchers. One study published by Nature magazine [14] shows that women are usually underrepresented in the peer review where only 20% of the reviewers are women. A similar study [15] shows that women and authors from emerging countries were underrepresented as editors and in peer reviewers. Because this problem is a focus area for the National Science Foundation (NSF), they developed an automated reviewer selection system that considers different demographic features when selecting reviewers [16]. The problem of bias in a peer review process is not limited to gender and race, but it can be seen from other angles, such as the geolocation of the reviewer. For example, a study in [17] shows that the US dominated the peer review process by 32.9% while its

 ISBN: 978-1-61208-765-8

publications represent only 25.4% of publications between 2013 and 2017.

III. PROBLEM FORMULATION

Given a set of conferences with n experts (i.e., authors) E such that $E = \{e_1, e_2, e_3,…., e_n\}$, our goal is to build a fair ranking system R that can be used to recommend experts to join a conference program committee. To achieve that, we will try to solve the following two problems. First, we need to quantify the set of expertise that each expert possesses. Hence, we define the set S that contains the expertise score for each expert in E such that $S = \{s_1, s_2, s_3,…., s_n\}$ where s_i is a scalar number represents the expertise score for expert e_i where $1 \leq i \leq n$. Second, we will define a demographic profile for each member of the set E with respect to a set of categorial features F with cardinality m such that $F = \{f_1, f_2, f_3,….,f_m\}$. Hence, for expert e_i, we have the set $D = \{D_1, D_2, D_3,…., D_n\}$ where $D_i = < d_{i1}, d_{i2}, d_{i3},…, d_{im} >$ that represents the demographic profile based on features in the set F. Finally, recommendation algorithms (FUNC) will be proposed based on S, D, or both that provide different ranking for the set E such that FUNC(S,D) $\rightarrow$ R where $R = \{R_1, R_2, R_3,…., R_p\}$where p is the number of ranking produced by FUNC.

IV. EXPERT MODELING

A. Expertise Profiling

Expertise profiling can be defined as a record that shows the proficiency of specific knowledge areas that an expert possesses [18]. It can be viewed as a vector of scores that shows the competency of each skill for that expert. In academia, there has been considerable interest in developing expert profiles due to the demand of having experts to review papers, participate in conference program committees or grant review panels, or finding talented individuals to join research teams. There have been different attempts to measure the amount of expertise that an expert has in academia as in [4][16][19]-[22]. One method that we propose here is to use the *h*-index as a metric to assess the scientific performance of a researcher. *h*-index was proposed by Hirsch in 2005 to measure the researcher's quality and productivity [22]. It is a robust single-number metric that uses the number of publications to indicate the quality of the researcher's output and the citation count to represent the quality of the expert's work. *h*-index scores are also employed by funding bodies and employers to determine funding, career decisions, promote and award committees [23][24]. Using a single score number to assess researcher expertise helps to rank those candidates and finally makes these decisions much easier [25]. It has been incorporated in many scholarly databases such as Google Scholar, Web of Science, Scopus, and Publish or Perish. In this research, we will use *h*-index as provided by Google Scholar as our metric to represent the expertise score for each researcher, as it tends to offer more excellent coverage and accuracy for computer scientists compared to other bibliometric databases [26].

B. Demographic Profiling

Many academic institutions and scholars realized the significance of including the demographic features in an expert profile for the reason that such data can be used in discrimination countermeasures, achieving fairness goals, complying with state and local regulation with respect to fair and diverse employment opportunity. In this research, we will represent the demographic profile using five important features that have been considered major sources of bias in the academic environments that are: gender [27][28], race [28][29], career stage [30], institution geolocation [17], institution ranking [17]. To build that profile, we will incorporate different techniques of feature predictions and web crawling to collect the attributes of the demographic profile as some features are explicitly included in researcher personal home pages while others might not be included due to privacy concerns, and hence prediction tools will be employed to predict such a feature.

To predict the gender and race, we will use NameSor software described in [13].The software uses a database of name information of more than 4 billion names [31] with the help of a novel machine-learning algorithm to provide a matching probability for the gender and race. One challenge is identified by [13] with respect to predicting gender with Chinese names. We manually validated any name that had a gender confidence probability of 0.6 or less, and we found that a gender matching accuracy of 80% with respect to Chinese names and 92% percent with respect to others, and we manually rectified any discrepancies. However, the accuracy was not as high when predicting race, specifically predicting African American as the system provides an accuracy of 15% by labeling many White scholars as African American. Hence, we manually verified every scholar labeled as African American by NamSor to correct any errors. Nevertheless, the software predicts other races with an acceptable accuracy of 75-80%. Once scholar gender and race have been predicted, we map it from categorical to binary values by using the concept of protected parameters where minorities are assigned a value of 1 and the majority will be penalized by having a value of 0. For example, gender will have these values (0 for male and 1 for female) while the race of White and Asian has 0, and other races will have a value of 1.

Career stage extracted from the Google Scholar (GS) page and mapped to binary values by having two classes (junior = 1, senior = 0). We define senior researchers as any researcher who has the academic rank of an associate professor, a senior lecturer, or above. Any academic rank lower than associate professor is considered a junior researcher. Geolocation is collected using the same method of the career stage. We map this to a developed country (Binary value: 0) and developing country (Binary value: 1) as per the last United Nations countries classification [32]. The last attribute is the affiliation rank, for which we used the TIMES computer science university ranking system [33] and mapped it to 0 if the university rank is less than the mean of TIMES computer science university ranking and 1 otherwise.

V. Scholar Recommendation

In the previous section, we demonstrated the expertise and demographic profiles that are considered the main inputs to our recommendation algorithms. Now, we are ready to describe our proposed scholar recommendation algorithms that address the issue of recommending an expert. We consider the case of recommending a researcher to join a team of conference program committee. We will address this problem from three different aspects: expertise, demographic diversity, and a balanced approach between expertise and diversity.

A. *Expertise Recommendation Approach*

As discussed above, we quantified the skills of scholars using *h*-index, where the higher the score is, the higher the expertise that such scholar has. To add a researcher to a team (e.g., conference program committee), we get the *h*-index of every author who published an article in that conference, extract GS *h*-index, and recommend the scholar that has the highest *h*-index. One advantage of this approach is that it maximizes the expertise in the process of the recommendation. However, it might lead to a systematic bias by favoring one demographic group over the other by failing to address the issue of demographic diversity. For example, the expertise-only approach does not consider the issue of the gender gap and the race gap. Hence we might end up with a team of the same race or gender. Another concern is by favoring highly cited researchers; it minimizes the opportunities for junior researchers to attend such research teams, which negatively impacts their chances to advance their careers. Additionally, most highly cited researchers are employed by the top rank universities, and this approach would less favor those researchers from lower-tier universities.

B. *Diverse Recommendation Approach*

The second algorithm utilizes the demographic profile for a scholar as the basis by which to recommend a researcher. In this approach, we calculate the diversity score for a researcher. The diversity score is simply the sum of the binary score for demographic features, where d_i is the diversity score for each feature and n is the number of demographic features in the profile (see (1)). For example, if the demographic profile for a researcher is (gender = woman, race= African American, Career Stage = professor, University rank = 2, Country = United States) then the corresponding diversity score is (1+1+0+0+0 = 2). We refer to this algorithm as (**DIV**) where the scholars will be ranked according to their diversity score and in descending order. In case two or more researchers have the same score, the algorithm will randomly pick one to be recommended.

$$\text{Score(DIV)} = \sum_{i=1}^{n} d_i \qquad (1)$$

C. *Hybrid Recommendation Approach*

The previous two approaches each have advantages and disadvantage. The first approach enhances the expertise of the team but fails to address the problem of forming a diverse team. The diversity approach solved that problem, but again it might cause an unacceptable drop in the expertise level of the team. Hence, a hybrid approach is introduced. That introduces a tuning parameter (α) to balance these approaches, as shown in (2):

$$\text{Score(H)} = [\alpha * \text{Score(DIV)}] - [(1-\alpha) * \text{Score (EXP)}] \qquad (2)$$

Our goal, in this case is minimizing the utility loss, which results from favoring the diversity over the expertise to the minimum, hence we will test different values for α. To make both scores comparable, we will measure the performance using score scaling such that Score (EXP) and Score (DIV) will be normalized so that both scores will get a value between (0 and 1). Equation 3 will be used in our normalization process

$$\text{Score(H)norm} = \frac{Score_i - \min(Score)}{\max(Score) - \min(Score)} \qquad (3)$$

VI. Experiment

A. *Dataset*

We will test our recommendation algorithms by recommending scholars to join an existing conference program for three of The Association for Computing Machinery (ACM) conferences that have high impact factors [34]. These conferences' Program Committees (PCs) were previously found to be less diverse than the set of accepted authors [35][36]. For the year 2017, we collected information about all the authors and PC members for SIG-CHI (The ACM Conference on Human Factors in Computing Systems), SIG-MOD (International Conference on Management of Data), and SIGCOMM (The ACM Conference on Data Communication). Using information on their Google Scholar page and home page, we collected the demographic information, discarding researchers in industry and that missing demographic information. The total profiles in our dataset are 1217 and can be seen in Table I.

TABLE I. DATASET DESCRPTION

Conference	PC members	Authors
SIGCHI17	213	436
SIGMOD17	130	290
SIGCOMM17	23	125

B. *Baseline and Metric*

For each dataset, we generate different K ranking, where K is the ranking cutoff, using our proposed algorithms and the following baseline:

 ISBN: 978-1-61208-765-8

Baseline: We used the Expertise Approach that selects candidates based on qualifications (*h-index*) only as our baseline (see Section V-A) for comparison.

Metric 1: Diversity gain based on mnDCG: We will compare the baseline with other approaches using normalized discounted cumulative gain (nDCG) by calculating the diversity gain at each rank. However, nDCG works with one feature at a time, hence, we had to modify the nDCG metric to support multi-feature ranking gain. Thus, we propose multidimensional normalized Discount Cumulative Gain (mnDCG) that can be calculated in three steps. First we calculate the DCG per feature as in (4) where n is the maximum rank and score (*f, i*) is the score for feature *f* for the candidate *i* in the expert demographic profile. Once DCG is calculated, then Ideal Discounted Cumulative Gain (IDCG), is calculated for each feature by ranking candidates in a descending order based on that feature. Now, nDCG for that feature can be calculated using (5). The process repeats itself for all features and the mnDCG is the average nDCG gain over all features as shown in (6).

$$DCG = \left(\sum_{i=1}^{n} \frac{2^{score(f,i)}}{\lg(1+i)}\right) \quad (4)$$

$$nDCG_f = DCG_f / IDCG_F \quad (5)$$

$$mnDCG = \frac{1}{k}\sum_{j=1}^{k} nDCG_f \quad (6)$$

Metric 2: F-Measure: We will use the F-measure as harmonic mean between the diversity and expertise gain.

VII. RESULTS AND DISCUSSIONS

Table II displays the result of evaluating our diversity approach against baseline and RAND recommendation algorithm which randomly picks candidates. We tested our proposed algorithms on each conference separately and measure the diversity gain by recommending top K experts for different values for K. We reported the diversity gain using mnDCG for each ranked set of candidates produced by our ranking algorithms. As Table II presents, the DIV algorithm always outperforms the other algorithms with respect to diversity. We also notice that the expertise algorithm produces the poorest diversity performance as compared to other algorithms, including random, indicating that it produces program committees that do not reflect the demographics of the community as a whole.

TABLE II. MNDCG DIVERSITY GAIN RANKING PRODUCED BY RAND, EXPERTISE, AND DIVERSITY ALGORITHMS

Conference	Rank@K	RAND	DIV	Expertise
SIGCHI17	50	0.222	0.617	0.113
	100	0.23	0.66	0.122
	Total (436)	0.639	0.847	0.602
SIGCOMM17	50	0.374	0.679	0.29
	100	0.494	0.804	0.523
	Total (125)	0.639	0.804	0.602
SIGMOD17	50	0.207	0.563	0.164
	100	0.312	0.66	0.227
	Total (290)	0.648	0.821	0.608

Nevertheless, promoting diversity comes at the cost of expertise. Hence, we tested our balanced approach presented in Section V-C that incorporates the results from the diversity algorithm (DIV) and the expertise algorithm using a linear tuning parameter (α). Table III shows these results averaged over the three conferences. We report the expertise saving to represent the amount of expertise retained after incorporating diversity, and the diversity gain relative to the baseline expertise algorithm. We use F-measure to combine the two diversity and expertise gains into a single metric. We report the result using α using steps of 0.1 to find the best value. α of 0 indicates the expertise only algorithm and α 1.0 indicates the diversity only algorithm. The highest F-measure is achieved when alpha is 0.4 indicating a 60% contribution from the expertise ranking and 40% from the diversity algorithm.

TABLE III. HYBRID ALGORITHM EVALUATION

α	Diversity Gain	Expertise Gain	F-Measure	Diversity Gain%	Expertise Saving %
0	0.603	1	0.752	0%	100%
0.1	0.642	0.998	0.781	6.60%	99.80%
0.2	0.659	0.993	0.792	9.40%	99.30%
0.3	0.69	0.975	0.808	14.50%	97.50%
0.4	0.731	0.922	**0.816**	**21.30%**	**92.20%**
0.5	0.784	0.829	0.806	30%	82.90%
0.6	0.813	0.771	0.792	35%	77.10%
0.7	0.829	0.671	0.742	37.70%	67.10%
0.8	0.832	0.609	0.703	38.10%	60.90%
0.9	0.832	0.608	0.703	38.10%	60.80%
1	0.824	0.554	0.662	36.70%	55.40%

To illustrate the effect of our approach, we provide the participation of members of underrepresented groups using the baseline recommendation algorithm versus our hybrid approach (with α = 0.4, rank@K = 50) in Figure 1. For example, the results show that our balanced, hybrid algorithm, has increased the females in SIGCOMM from 22% to 44%, the developing countries in SIGMOD from 16% to 56%, and racial minorities from 4% to 14% in SIGCHI with 92% expertise saving.

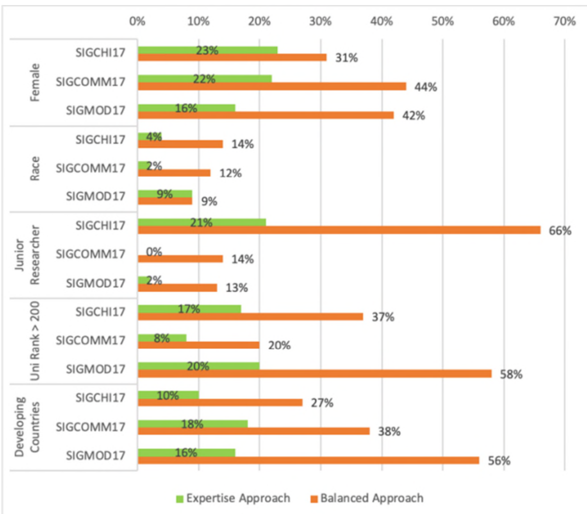

Figure 1. Demographic gain with average expertise loss of 7.8% @ K = 50 using α = 0.4.

We also test our hybrid approach by recommending the same PC size from a pool of the real PC and conference authors and compare it to the demographic distributions of the real PC, as shown in Figure 2. Our algorithm increased the representation of all demographic groups on average across the three conferences. The average expertise loss, as measured by the nDCG on the *h*-index, was 1.3%, a small penalty to pay for increased diversity.

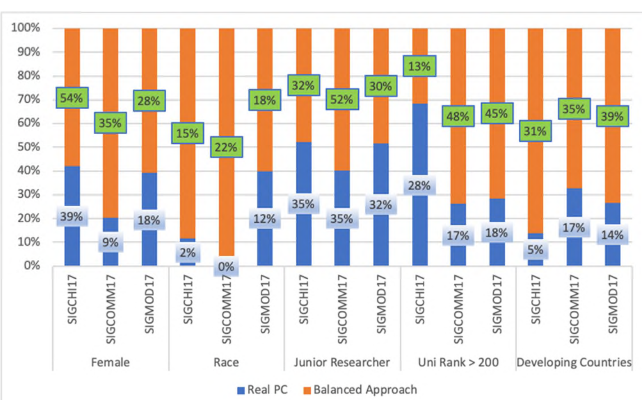

Figure 2. Demographic difference between real PC and balanced approach [α = 0.4].

VIII. CONCLUSION

The paper presents an approach to incorporate demographic fairness in expert recommendations in academia. It also introduces a more comprehensive way to represent demographics in researcher profiles in order to achieve fairness, increase demographic diversity, and ensure that members of underrepresented demographic groups have access to career opportunities. Our profiles include five attributes that have been shown to be sources of explicit or implicit bias in academia, i.e., gender, race, career stage, academic rank, and affiliation geolocation. We use these demographic features within an expert recommender system in academia. The paper presents and evaluates three scholar recommendation approaches: 1) the expertise model; 2) a new diversity model; and 3) a balanced approach between that balances diversity gains against loss of expertise. We consider a specific example of expert recommendation in academia that is recommending researchers to join a conference program committee. We created a dataset of 1217 researcher profiles from the three top ACM conferences for 2017. We evaluate our algorithms using a modified nDCG metric, mnDCG, that measures gain across multiple dimensions. Our evaluation shows our diversity approach provides a better diversity gain; however, this comes with the cost of expertise. Hence, we developed a hybrid recommender system that incorporates linear optimization through a tuning parameter. Our results show that the best parameter value for the three conferences studies is approximately 0.4, i.e., 40% weight to the diversity recommendation and 60% weight to the expertise recommendation.

In the future, we will extend the demographic profile design to contain continuous values to provide a wide range of demographic groups for the same attribute. We will apply these new profiles to fair group formation algorithms [36]. We intend to assign different weights to each demographic feature based on different mechanisms and study whether this leads to a better demographic representation. Also, we plan to study the demographic composition of different academic conferences in other domains.

REFERENCES

[1] M. Winans. *10 Key Marketing Trends for 2017 Customer Expectations and Ideas for Exceeding Customer Expectations*. [Online]. Available from: URL:ftp://ftp.www.ibm.com/software/in/pdf/10_Key_Marketing_Trends_for_2017.pdf [retrieved: October, 2020]

[2] W. Reinhardt, B. Schmidt, P. Sloep, and H. Drachsler, "Knowledge worker roles and actions—results of two empirical studies," Knowledge and Process Management, vol. 18, pp.150-174, Jul. 2011.

[3] K. Balog, Y. Fang, M. de Rijke, P. Serdyukov, and L. Si, "Expertise retrieval," Foundations and Trends® in Information Retrieval, vol. 6, pp.2–3, Feb 2012.

[4] D. Pedreshi, S. Ruggieri, and F. Turini, "Discriminationaware data mining," Proc. of KDD, ACM Press, Aug. 2008, pp. 560–568.

[5] J. Tang et al., “Arnetminer: Extraction and mining of academic social networks,” Proc. of KDD, ACM Press, Aug. 2008, pp. 990-998.

[6] M. Cochran-Smith and K. Zeichner, Studying teacher education: The report of the AERA panel on research and teacher education, Routledge, 2009.

[7] J. Michael, “40000 namen, anredebestimmung anhand des vornamens,” C'T, 2007, pp. 182-183.

[8] I. Perez. *Gender-guesser*. [Online]. Available from: https://pypi.python.org/pypi/genderguesser [retrieved: October, 2020]

[9] M. Vanetta. *Gender Detector* [Online]. Available from: https://github.com/malev/genderdetector [retrieved: October, 2020]

[10] R. Knowles, J. Carroll, and M. Dredze, “Demographer: Extremely simple name demographics,” Proceedings of the First Workshop on NLP and Computational Social Science, Nov. 2016, pp. 108-113.

[11] J. Ye et al., “Nationality classification using name embeddings,” Proc. of the 2017 ACM on Conference on Information and Knowledge Management, Nov. 2017, pp. 1897-1906.

[12] C. Strømgren. *Genderize.io*. [Online]. Available from: https://genderize.io/ [retrieved: October, 2020]

[13] S. Mattauch, K. Lohmann, F. Hannig, D. Lohmann, and J. Teich,” A bibliometric approach fo detecting the gender gap in computer science,” Communications of the ACM, vol. 63, pp. 74-80. Apr. 2020.

[14] J. Lerback and B. Hanson, “Journals invite too few women to referee,” Nature, vol. 541, pp. 455-457, Jan. 2017, doi:10.1038/541455a

[15] H. Yin, B. Cui, and Y. Huang, ” Finding a wise group of experts in social networks,” Proc. of International Conference on Advanced Data Mining and Applications, Dec. 2011, pp.381-394. Springer, Berlin, Heidelberg.

[16] S. Hettich and M. Pazzani, “Mining for proposal reviewers: Lessons learned at the national science foundation,” Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 2006, pp. 862-871.

[17] Publons. *Global State Of Peer Review*. [Online]. Available from https://publons.com/static//Publons-Global-State-Of-Peer-Review-2018.pdf [retrieved: October, 2020]

[18] K. Balog and M. De Rijke, “Determining expert profiles (with an application to expert finding),” Ijcai, vol. 7, pp. 2657-2662, Jan. 2007.

[19] K. Chandrasekaran, S. Gauch, P. Lakkaraju, and H. Luong, “Concept-based document recommendations for citeseer authors,” The International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Jul. 2008, pp. 83-92.

[20] A. Pudhiyaveetil, S. Gauch, H. Luong, and J. Eno, “Conceptual recommender system for CiteSeerX,” Proceedings of the third ACM conference on Recommender systems, ACM, Oct. 2009, pp. 241-244.

[21] B. Sateli, F. Löffler, B. König-Ries, and R. Witte, “ScholarLens: Extracting competences from research publications for the automatic generation of semantic user profiles,”PeerJ Computer Science, vol. 3, e121, Jul. 2017.

[22] J. Hirsch, “An index to quantify an individual's scientific research output,” Proceedings of the National academy of Sciences, vol. 102, pp.16569-16572, Nov. 2005.

[23] *Measuring Research Success:H-index Scores in Science*. [Online]. Available from: fromhttps://conductscience.com/measuring-research-success-h-index-scores-in-science/ [retrieved: October, 2020]

[24] L. Tang and G. Hu, “Evaluation woes: Metrics can help beat bias,” Nature, vol. 559, p. 331, Jul. 2018.

[25] W. Schreiber and D. Giustini, “Measuring scientific impact with the h-Index: a primer for pathologists,” American journal of clinical pathology, vol. 151, pp. 286-291, Feb. 2019.

[26] J. Bar-Ilan, “Which h-index?A comparison of WoS, Scopus and Google Scholar,” Scientometrics, vol. 74, pp. 257-271, Feb. 2008.

[27] L. Bornmann and H. Daniel, “Reliability, fairness and predictive validity of committee peer review,” BIF Futura, vol. 19, pp. 7-19, 2004.

[28] M. McConner, “Black women in leadership: An assessment of the gender inequality and racism that exists among black women leaders in higher education,” The Journal Higher Education Management, vol. 29, pp. 78-87, 2014.

[29] L. Perna, D. Gerald, E. Baum, and J. Milem, “The status of equity for black faculty and administrators in public higher education in the south,” Research in Higher Education, vol. 48, pp. 193-228.

[30] S. Gallo, L. Thompson, K. Schmaling, and S. Glisson, “Participation and motivations of grant peer reviewers: a comprehensive survey of the biomedical research community,” Science and engineering ethics, vol. 26, pp. 761-782, Apr. 2020.

[31] D. Menéndez, “Damegender: writing and comparing gender detection tools,” EasyChair, Apr. 2020.

[32] United Nations. *UN World Economic Situation and Prospects (WESP) Report*. [Online]. Available from: https://www.un.org/development/desa/dpad/publication/world-economic-situation-and-prospects-2019 [retrieved: October, 2020]

[33] Times Higher Education. *World University Rankings 2019 by subject: computer science* [Online]. Available from: https://www.timeshighereducation.com/world-university-rankings/2019/subject-ranking/computer-science#!/page/0/length/25/sort_by/rank/sort_order/asc/cols/stats [retrieved: October, 2020]

[34] *Conference Rank* [Online]. Available from: https://cn.aminer.org/ranks/conf [retrieved: October, 2020]

[35] O. Salman, S. Gauch, M. Alqahatni, and M. Ibrahim, “The Demographic Gap in Conference Program Committees,” Proc. of the 17th International Conference on Applied Computing (AC 2020), Nov. 18-20, 2020 (in press).

[36] M. Alqahatni, S. Gauch, O. Salman, M. Ibrahim, and R. Alsaffar, “Diverse Group Formation Based on Multiple Demographic Features,” Proc. of the 12th International Conference on Knowledge and Discovering in Information Retrieval (KDIR 2020), Nov. 2-4, 2020.

 ISBN: 978-1-61208-765-8