Incorporating Diversity in Academic Expert Recommendation

Omar Salman, Susan Gauch, Mohammed Alqahatani, Mohammed Ibrahim, Reem Alsaffar

Presented by:
Omar Salman
Computer Science and Computer Engineering Department
University of Arkansas, Fayetteville, AR, USA
Email: oasalman@uark.edu

Presenter’s Bio’s

Omar is a Ph.D. candidate at the computer science and computer Engineering Department (CSCE) at the University of Arkansas. He is part of Text Analysis Research Lab led by Dr. Susan Gauch. His main research interests include: Algorithmic Fairness, Expert Recommendations, Information Retrieval, and Natural Language Processing.

Introduction

- We witness an exponential growth of knowledge
- In 2016, 90 % of available data were not available two years prior.
- Modern economies shifted to knowledge-based economies.
- Intellectual capabilities and expertise determine the individual values in an enterprise or a societies.
- Measuring the level of expertise is a challenge.
- Expert recommender system is a solution

Introduction

- Expert Recommendation System types :
  - Manual Systems
    - Individual inputs his or her own skills
    - Example : Yellow pages
    - Issues: manual update and accuracy problems

  - Automated System:
    - Intelligent Systems that extract skills from documents (e.g. Internal email communications)
    - Similar to search engine
    - Example: P@noptic expert System
    - Issues:
      1. Rely on textual representation of a document
      2. No standard terms to describe skills

Expert Recommendation in Academia

- Expert recommendation is not limited to industry
- In academia, it has been used to:

  - Hiring and recruiting process.

  -Recommend experts to evaluate patents.

  - Identifying reviewers for scientific conferences.

  - Assembling a conference program committee.

Expert Recommendation Challenges in Academia

- Accuracy
- Direct and indirect bias.
- Bias from ML algorithms.
- Lack of opportunity for junior researchers
- Not addressing fairness and diversity issues.
- Bias based on gender, race, and location is well documented in academia.

How to Address These Challenges

- Develop a unified representation (profile) for researchers that quantifies skills and demographic of a researcher.
- Include key demographic and socioeconomic information about researchers to ensure fair and diverse representation.
- Investigate recommendation algorithms to eliminate bias sources
- Develop accurate and fair recommendation algorithms that recommend diverse researchers.

Our Research

- **Problem** :Recommend experts to join a conference program committee.
- **Techniques**: Machine Learning, Expertise Retrieval System, and Information Retrieval.
- **Goal**: Provide efficient, fair, and diverse group representation

Research Plan

- We will have two goals to achieve that:

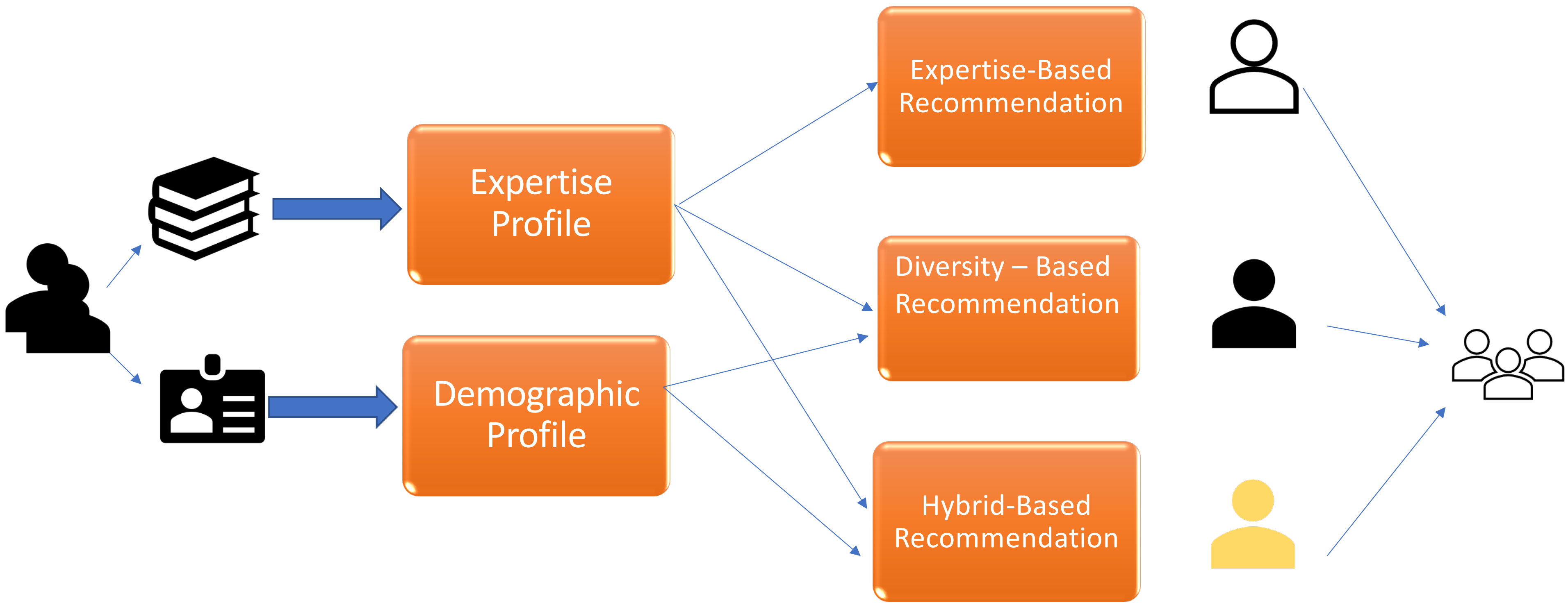
  Goal 1: Modeling a researcher: by modeling the expertise and demographic features of a researcher.

  Goal 2: Design algorithms to recommend researchers to a program committee based on the profile developed in Goal 1

Contributions

- 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.

Expert Recommendation - A Proposed System



Expertise Profile

- Expertise profiling can be defined as a record that shows the proficiency of specific knowledge areas that an expert possesses
- Challenge: How to describe skills in academia?
- We use the *h*-index as a metric to quantify the skills of a researcher.
- *h*-index was proposed by Hirsch in 2005 to measure the researcher's quality and productivity.
- *h*-index scores are also employed by funding bodies and employers to determine funding, career decisions, promote and award committees.
- Using a single score number to assess researcher expertise helps to rank those candidates and finally makes these decisions much easier.

Expertise Profile

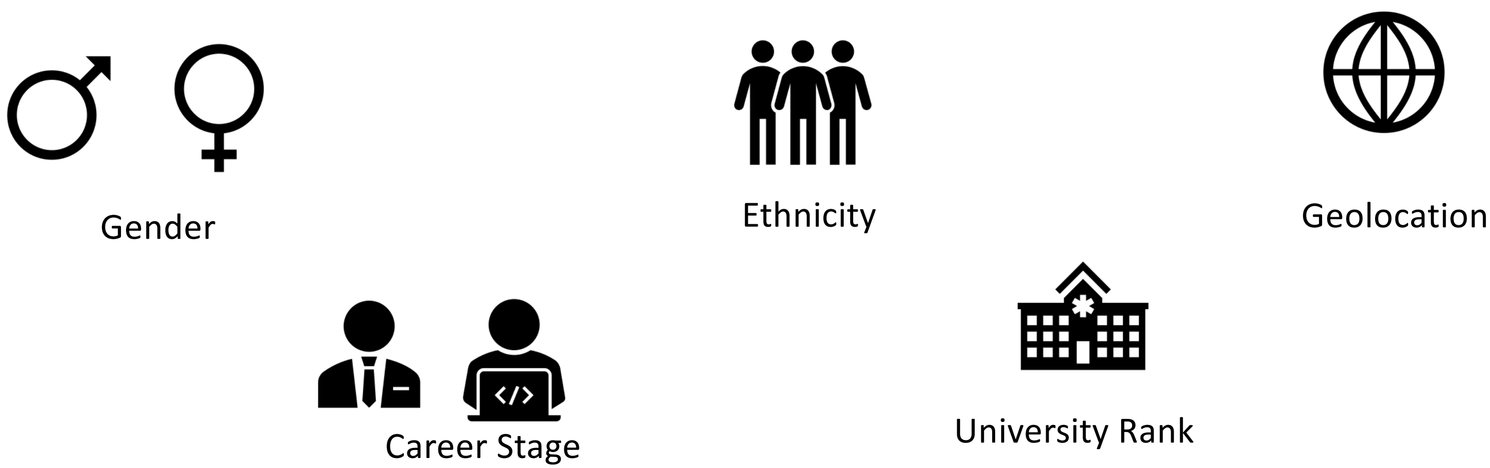
Cited by VIEW ALL

	All	Since 2015
Citations	7034	1973
h-index	34	18
i10-index	66	26

- Different scholarly databases (Google Scholar, Web of Science, Scopus, and Publish or Perish), different *h-index. Which one to consider?*
- Google Scholar appears to offer more excellent coverage and accuracy for computer scientists compared to other bibliometric databases as indicated by research.

Demographic Profiler

- Model the demographic information of a researcher
- Challenge: Privacy Concerns?
- What features to collect?



How to Predict Gender and Race

- We will use *NamSor* to predict gender.
- A database of more than more than 4 billion names.
- It uses novel machine-learning algorithm to provide a matching probability for the gender and race.
- Query a first and last name and return the gender with a confidence based on the distribution of that name across female and male.
- We will accept a confidence of 0.6 or more as a gender accuracy. Other case will be reviewed manually
- Our Validation: Accuracy is 80% with respect to Chinese names and 92% percent with respect to others.
- We manually rectified any discrepancies.

Validation

- 15% Accuracy to predict African race.
- We manually rectified any discrepancies.
- Nevertheless, the software predicts other races with an acceptable accuracy of 75-80%


```
"raceEthnicityAlt": "A",
"raceEthnicity": "W_NL",
"score": 15.763382065030502,
"raceEthnicitiesTop": [
  "W_NL",
  "A",
  "HL",
  "B_NL"

{
  "id": null,
  "firstName": "Susan",
  "lastName": "Gauch",
  "likelyGender": "female",
  "genderScale": 0.9746671528639042,
  "score": 22.31465088366339,
  "probabilityCalibrated": 0.9873335764319521
}
```

How to Get Other Demographic Parameters

Geolocation → Affiliated University Geolocation → Google Scholar

Career Stage → Google Scholar

University Rank → THE WORLD UNIVERSITY RANKINGS

Feature	Value
Gender	Female
Ethnicity	Black
Career Stage	Junior Researcher
Nationality	United States (Developed)
h-index	10
University Rank	129

Adding an Expert to A Group

- Task: Recommend researchers to join a group.
- Three approaches to achieve that:

  a.) Expertise Recommendation Approach.

  b.) Diverse Recommendation Approach.

  c.) Hybrid Recommendation Approach.

Approach 1 – Expertise Recommendation Approach

- To add researchers to a team (e.g., conference program committee):
  - Get the h-index of every author who published an article in that conference.
  - Extract Google Scholar(GS) h-index.
  - Rank the scholar with respect to h-index score.
  - Recommend the top ranked experts and according to the required size of the conference PC.
  - Advantage: Maximizes the expertise in the process of the recommendation.

Expertise Recommendation Approach Disadvantages

- This approach has several disadvantages:
  - Systematic bias: 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.
  - Less opportunities for junior researchers by favoring highly cited researchers.
  - High h-index researchers are employed by the top rank universities, and this approach would less favor those researchers from lower-tier universities.

Approach 2 – Diversity Approach

- A social science approach that addresses social inequality and bias.
- Protected parameters that are those demographic information that should not bias against.
- Protected parameters can be defined by the law or by the environment.
- We assume that protected parameters vary from environment to another. For example, the protected value for gender is not STEM education is not the same as in nursing.
- We will model the value for the protected parameter as binary variable that can be either 1 or 0.

Approach 2 – A Diversity Approach(DIV)

- Binary profile will be calculated for each expert.
- Experts are ranked according to the sum of their demographic features from equation 1 in a descending order.
- If two or more experts have the same diversity score, then the one that has the highest h-index will be ranked higher.
- Recommend the top ranked experts and according to the required size of the conference PC.

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

Protected Parameters

Group	X1	X2
Gender	Female	Male
Ethnicity	Non-white	White and Asian
Geo-Location	Developing countries	Developed countries
Career Stage	Junior researcher	Senior researcher
Institution Rank	> = 563 (mean rank)	< mean rank (563)

$R_{dem} < 0,1,1,1,1,1,0 >$

Researcher demographic Profile

Approach 3 – Hybrid Recommendation Approach

- Approach 1 enhances the expertise of the team but fails to address the problem of forming a diverse team
- The diversity approach solves that problem, but again it might cause a drop in the expertise level of a team .
- We introduce a hybrid (fair) approach that considers linear optimization to achieve a balance between the two approaches.

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

Approach 3 – Hybrid Recommendation Approach

- We will introduce different values for α and compare the performance at each step.
- To make the two scores comparable, we will normalize using min-max normalization.

$$Score(i)_{norm} = \frac{Score_i - \min(Score)}{\max(Score) - \min(Score)}$$

Evaluation - Dataset

- PC and author profiles of three top ACM conferences for the year of 2017 (ACM SIGCHI, SIGCMOD, SIGCOMM) were collected.
- Expertise and demographic profiles were built by extracting the data from Google Scholar and personal homepages of researchers.
- Only academic profiles are kept (Industry profiles were excluded).
- Total profiles are 1217.

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

Baseline and Metric

- Generate different K ranking, where K is the ranking cutoff, using our proposed algorithms and the following baseline:
- **Baseline:** We used the Expertise approach that selects candidates based on qualifications (*h-index*) only as our baseline.
- **Metric 1: Diversity gain based on mnDCG:**

  - Modified nDCG that support multiple features simultaneously.
  - Three steps calculation:

    Step 1: Calculate the Discounted Cumulative Gain (DCG) per feature as per equation 4

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

score(*f,i*) is the score for feature *f* for the candidate *i* in the expert demographic profile.

Baseline and Metric

Step 2: Ideal Discounted Cumulative Gain (IDCG) is calculated for each feature by ranking candidates in a descending order based on that feature (f) as in (5).

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

- **Step 3:** The process repeats itself for all features and the *mnDCG* is the average nDCG gain over all features as shown in (6).

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

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

Experiment

- For each dataset (i.e. conference), we recommend researchers from the authors' pool in each dataset to join a conference PC.
- We tested recommending different number of researchers by recommending (50, 100, and all authors in that conference) to join the PC.
- We compare the diverse recommendation approach to expertise recommendation and random recommendation approaches.
- We reported the *mnDCG* (the diversity gain) for each algorithm with the corresponding PC size.

Diversity Approach Evaluation

- the DIV algorithm always outperforms the other algorithms with respect to diversity.
- The expertise recommendation produces the poorest diversity performance as compared to other algorithms, including random.
- This confirms that considering expertise alone produces program committees that do not reflect the demographics of the community as a whole.

TABBLE 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

TABBLE 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%

Hybrid Approach Evaluation

- 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 combine the two diversity and expertise gains into a single metric.
- We report the result using α using steps of 0.1 , where 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.

Hybrid Approach Evaluation

- The hybrid approach outperforms the baseline (expertise) by increasing the representation of all underrepresented groups with a minimal expertise loss.

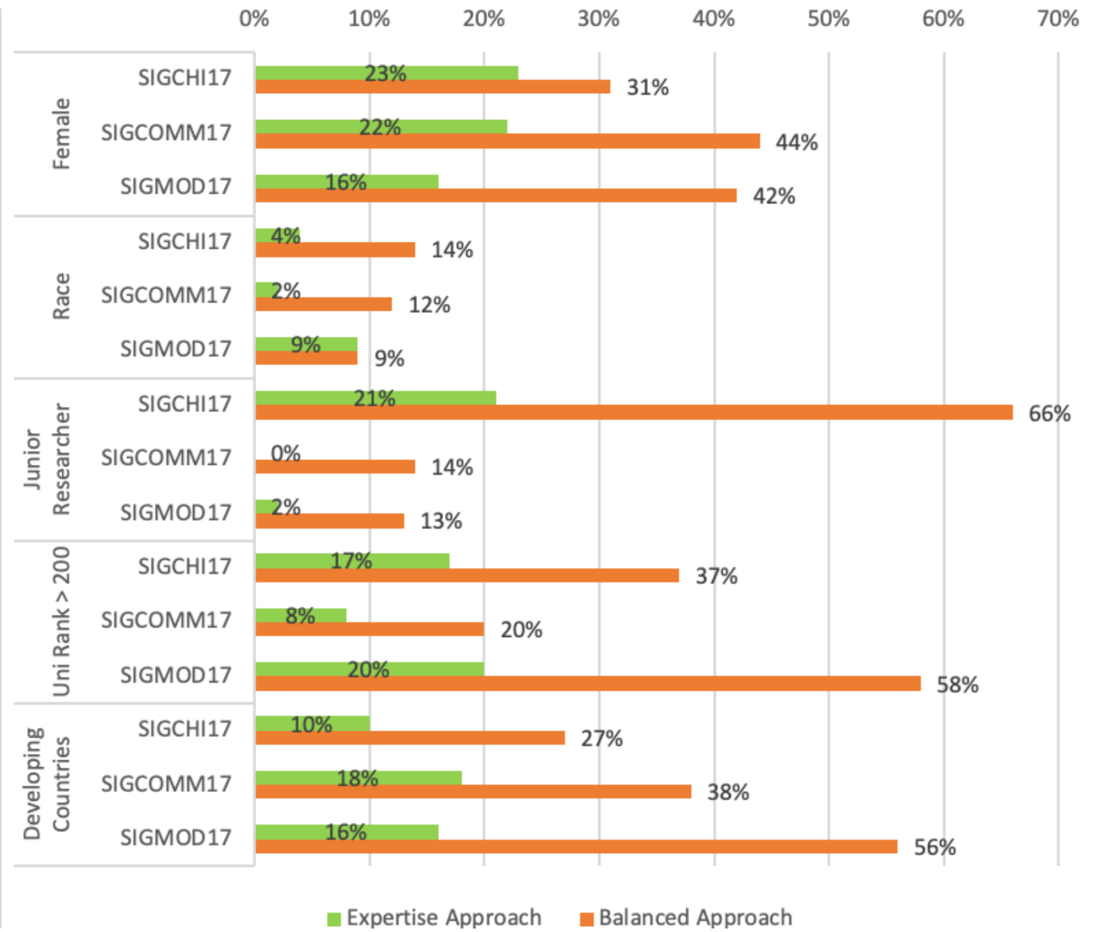


Figure 1 Demographic Gain with Average Expertise Loss of 7.8% When Recommending 50 researchers using α = 0.4

Figure 2. Demographic Difference Between Real PC and Balanced Approach [α = 0.4]

Hybrid Approach Evaluation

- We recommend 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 DCG on the *h*-index, was 1.3%, a small penalty to pay for increased diversity.

Conclusion

- We present an approach to incorporate demographic fairness in expert recommendations in academia .
- We introduce 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
- We evaluate three scholar recommendation approaches: 1) the expertise model; 2) a new diversity model; and 3) and a balanced approach between that balances diversity gains against loss of expertise.
- We created a dataset of 1217 researcher profiles from the three top ACM conferences for 2017

Conclusion

- We consider a specific example of expert recommendation in academia that is recommending researchers to join a conference program committee .
- We evaluate our algorithms using a modified nDCG metric, *mnDCG*, that measures gain across multiple dimensions.
- 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.

Future Work

- 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 .
- We intend to assign different weights to each demographic feature based on different mechanisms and study whether this leads to a better demographic representation.
- We plan to study the demographic composition of different academic conferences in other domains.