Comparing Fair Ranking Metrics

Amifa Raj
People and Information Research Team, Boise State University

Connor Wood
People and Information Research Team, Boise State University

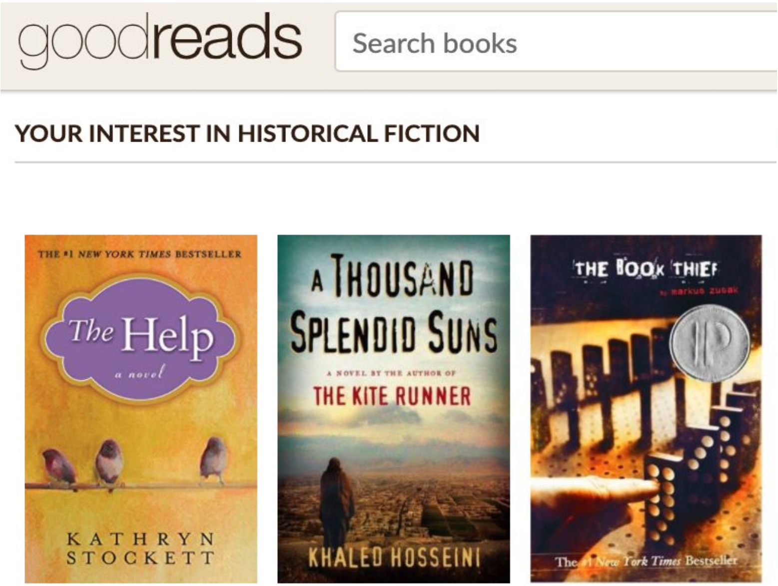
Ananda Montoly
Smith College

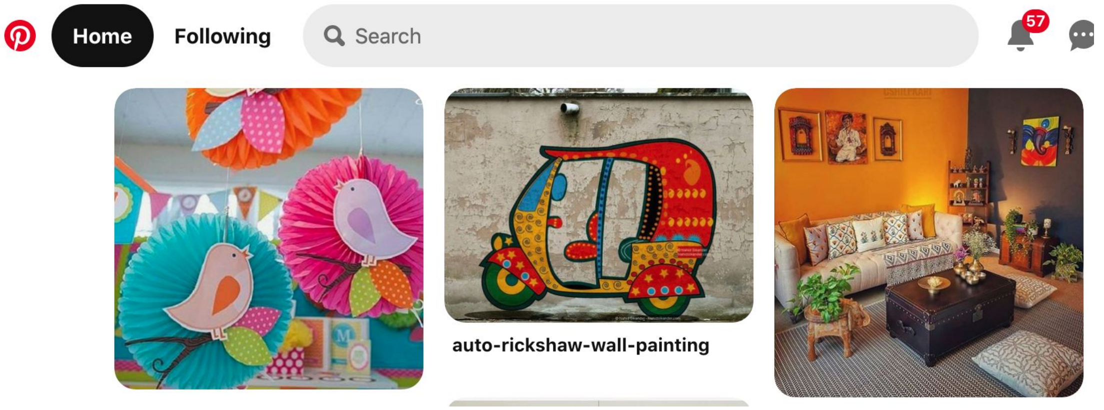
Michael D. Ekstrand
People and Information Research Team, Boise State University

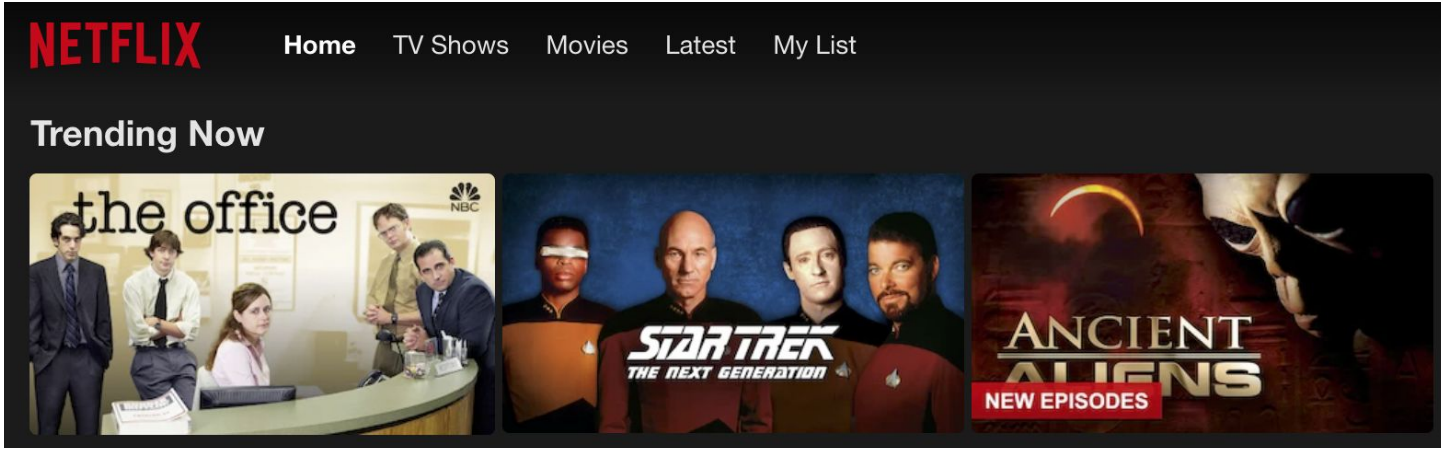
People and Information Research Team

BOISE STATE UNIVERSITY

grant IIS 17-51278

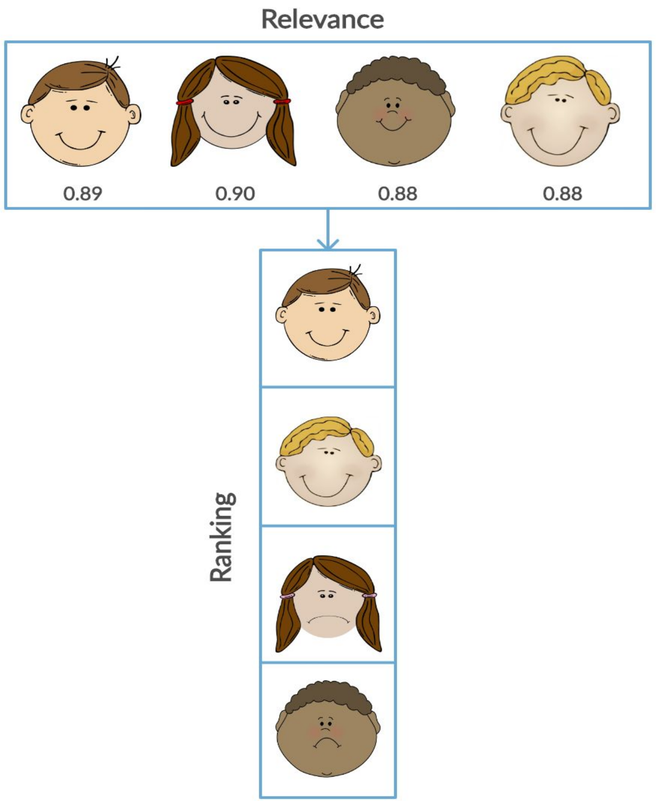

goodreads
Search books
YOUR INTEREST IN HISTORICAL FICTION
THE #1 NEW YORK TIMES BESTSELLER
The Help
a novel
KATHRYN STOCKETT
A THOUSAND SPLENDID SUNS
A NOVEL BY THE AUTHOR OF
THE KITE RUNNER
KHALED HOSSEINI
THE BOOK THIEF
The #1 New York Times Bestseller


Home
Following
Search
57
auto-rickshaw-wall-painting


NETFLIX
Home
TV Shows
Movies
Latest
My List
Trending Now
the office
NBC
STAR TREK
THE NEXT GENERATION
ANCIENT ALIENS
NEW EPISODES

Motivation

Bias in ranking



Disparate exposure

Several metrics to measure (un)fairness

Fair Ranking Metrics Resources

SSDBM '17

Measuring Fairness in Ranked Output
Yang et. al.

PreFd

SIGIR '18

Equity of Attention: Amortizing Individual Fairness in Rankings
Biega et. al.

IAA

KDD '18

Fairness of Exposure in Rankings
Singh et. al.

DP, DIR, DTR

WWW '19

Quantifying the Impact of User Attention Fair Group Representation in Ranked List
Sapienzynski et. al.

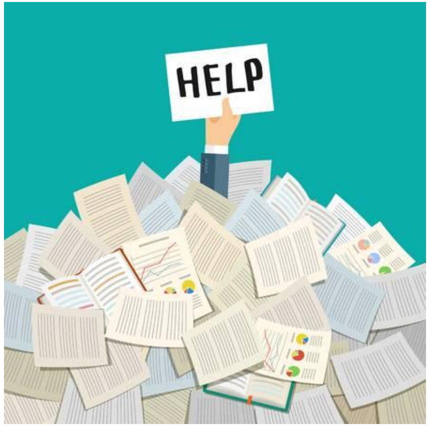
AWRF

CIKM '20

Evaluating Stochastic Rankings with Expected Exposure
Diaz et. al.

EER, EEL, EED

Why the problem is a problem!



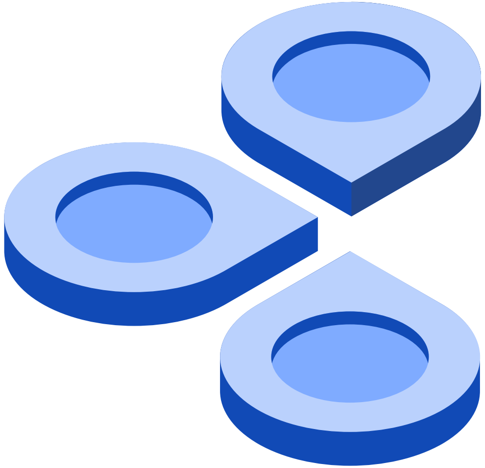
Several metrics to measure (un)fairness

Finding the suitable metric

Differences among the metrics

Contribution

Describe and compare exposure and rank-fairness metrics in **unified** framework



Identify gaps between their original presentation and the practicalities of applying them to recommender systems

Direct comparison of their outcomes with the same data and experimental setting

Fairness Positioning

Provider fairness

Group fairness

Fairness Definition

PreFd
AWRF

Item position should not be affected by membership *prioritizing parity at the top*

AWRF

Weighted ranking position

IAA
D*
EE*

Exposure/attention should be proportional to **relevance**

Target exposure is derived from relevance

EE*

D*: DP, DIR, DTR; EE*: EEL, EER, EED

Classification of Fair Ranking Metrics

Single-List Metric

PreFd
AWRF

Distribution and Sequence Metrics

IAA
D*
EE*

Summary of Fair Ranking Metrics

Metric(s)	Goal	Weighting	Relevance	Binary
PreFd	Each prefix representative of whole ranking	❌	❌	Dep on d
AWRF	Weighted representation matches population	Geometric	❌	❌
DP	Exposure equal across groups	Logarithmic	❌	✓
DTR	Exposure proportional to relevance	Logarithmic	✓	✓
DIR	Discounted gain proportional to relevance	Logarithmic	✓	✓
IAA	Exposure proportional to predicted relevance	Geometric	Predicted	❌
EEL, EER	Exposure matches ideal (from relevance)	Cascade, Geom	✓	❌
EED	Exposure well-distributed	Cascade, Geom	❌	❌

Experimental Setup

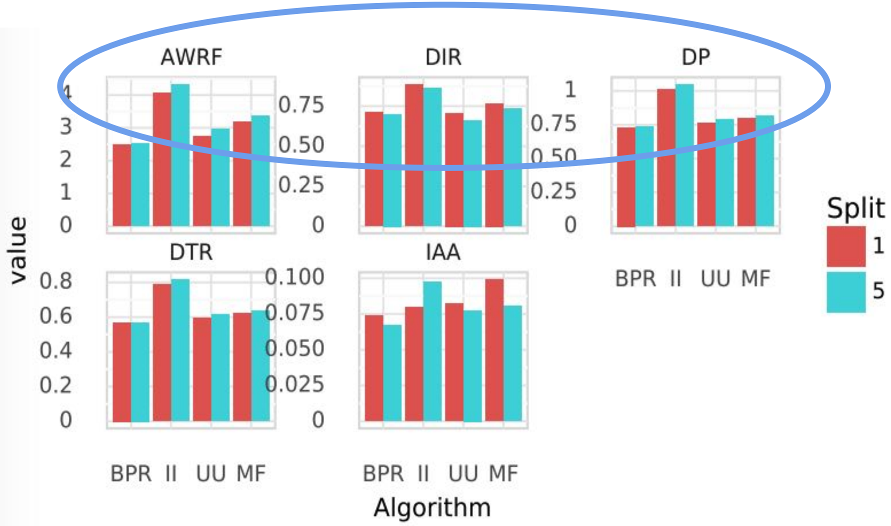
- Dataset
  - GoodReads book data (implicit feedback)
- Sensitive Attribute
  - Gender of author
- Recommendation Algorithms
  - user-based CF (UU)
  - item based CF (II)
  - matrix factorization (MF) and
  - Bayesian Personalized Ranking (BPR)

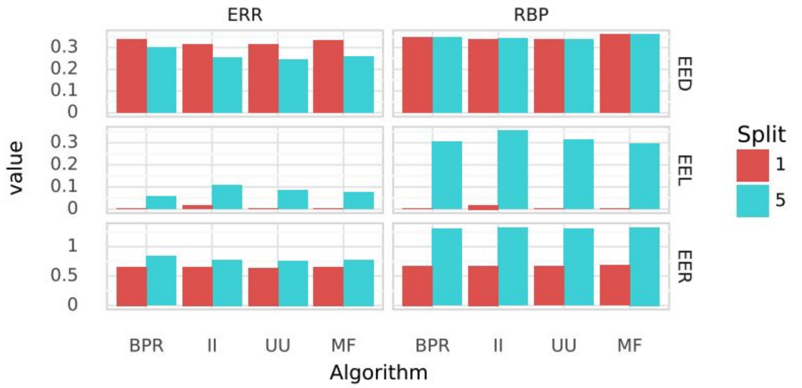
Two samples of 5000 users

- Split 1: each user rated at least **5** books, **1** held out
- Split 5: each user rated at least **10** books; **5** held out

Comparative Analysis

- Algorithms did not show significant differences on most metrics (exception: II)
- Size of relevance set has more effect on EE* than the choice of user model
- No clear agreement





Conclusion

Summary

- Unified the metrics under one framework
- Metrics are surprisingly similar
- Direct comparison did not provide conclusive result
- Missing pieces in implementing on real data

Future Work

- Missing labels
- Missing Relevance
- Sensitivity analysis
- More metrics
- More datasets

Request for Feedback!!!

Fair Ranking Metrics

PreFd Prefix Fairness Family	**AWRF** Attention-Weighted Rank Fairness	**IAA** Inequity of Amortized Attention
DP Demographic Parity	**DTR** Disparate Treatment Ratio	**DIR** Disparate Impact Ratio
EEL Expected Exposure Loss	**EER** Expected Exposure Relevance	**EED** Expected Exposure Disparity