BiasMeter: On Measuring Bias in Online Information

Panagiotis Papadakos, Irini Fundulaki, Giorgos Flouris
FORTH-ICS – Information Systems Laboratory

In collaboration with Evaggelia Pitoura and Panayiotis Tsaparas
Department of Computer Science and Engineering – University of Ioannina

MOTIVATION

Our diverse information needs are satisfied by online services

... but non-transparent algorithms return information that is ‘believed’ to be useful
... guiding our decisions in various domains

... the convenience and effectiveness of such services has limited our information seeking abilities and made us overly dependent on them

We live in echo chambers and filter bubbles due to personalization

TYPES OF BIAS

We consider bias in terms of the results of a service regarding a topic (e.g. politics)

CONTENT

• User independent
• Over a number of differentiating attributes political party, against / in favor, artists / scientists

A Web Search engine that returns results in favor of a specific political party

USER

• Different users
• Over a number of protected attributes race, sex, nationality, religion

A job recommendation engine that returns lowered paid results for women

BIASMETRER

DEFINITION 1 (INDIVIDUAL USER BIAS)
An online information provider (OIP) is individual user unbiased if for any pair of users \( u_1, u_2 \) it holds:

\[
D_R(R_{u_1}, R_{u_2}) \leq D_u(u_1, u_2)
\]

where \( R_{u_1} \) and \( R_{u_2} \) are the result lists received by \( u_1, u_2 \) resp. and \( D_R \) and \( D_u \) appropriate rankings and users distance functions

DEFINITION 2 (GROUP USER BIAS)
An OIP is group user unbiased if it holds:

\[
D_R(R_P, R_Q) \leq \varepsilon
\]

where \( R_P \) is the union of the result lists seen by the members of the protected attribute and \( R_Q \) is the union of the result lists seen by the members of the non protected attribute for some small \( \varepsilon \geq 0 \)

DEFINITION 3 (CONTENT BIAS)
An OIP is group user unbiased if it holds:

\[
D_R(R_T, R_T) \leq \varepsilon
\]

where \( R_T \) is the “ideal unbiased ranking” for some small \( \varepsilon \geq 0 \)

METRICS

RESEARCH CHALLENGES

1. Ground Truth
   - human-in-the-loop (crowd-sourcing)
   - comparative evaluation of various services

2. Multi-faceted and concrete Bias metrics
   - take into account the correlation of attributes

3. Engineering and Technical Challenges
   - generation of huge samples of user profiles
   - data mining / data integration / machine learning
   - knowledge representation / entity detection

4. Auditing Algorithms
   - access to the internals of services
   - co-operation with law and policy makers

Contact: Panagiotis Papadakos (papadako@ics.forth.gr), Irini Fundulaki (fundul@ics.forth.gr)