FairNet: A Genetic Algorithm to Reduce Marginalization in Social Networks

What is Marginalisation?

Discrimination in the literature: In a network, *algorithmic discrimination* focuses on the possible discrimination arising from an individual's network position, whereas the Data Mining field proposes that in a fair dataset *similar individuals should be treated similarly.*

Discrimination as marginalisation: Marginalisation is defined as *the act of relegating someone or something to an unimportant position*, and can indeed occur between different groups (e.g., a non-white person marginalised by white people) or inside of the same group (e.g., a white person marginalised by other white people). In a fair network, all nodes should be surrounded by a group of peers that manifests a similar distribution with respect to an attribute. Additionally, such distribution should be representative of the label distribution in the whole system. A proportionally-different distribution implies some sort of marginalisation against the node – either by nodes with different labels or by those with the same label.

Quantifying

We introduce **Individual Marginalisation Score** (IMS), a measure that takes into account the attribute distribution in the node's neighbourhood and compares it to the distribution in the whole network. IMS ranges in [-1, 1] and describes:

- marginalisation perpetrated by nodes with the same attribute for IMS < 0;
- marginalisation perpetrated by nodes with a different attribute for IMS > 0;
- no marginalisation for IMS $= 0$.

A node can be considered marginalised if its absolute IMS is beyond a fixed threshold.

The number of marginalised nodes can quantify marginalisation at the macro-scale level scale. We also the introduce the **System Marginalisation Score** (SMS), which captures the average marginalisation for all nodes, regardless of the sign.

Reducing

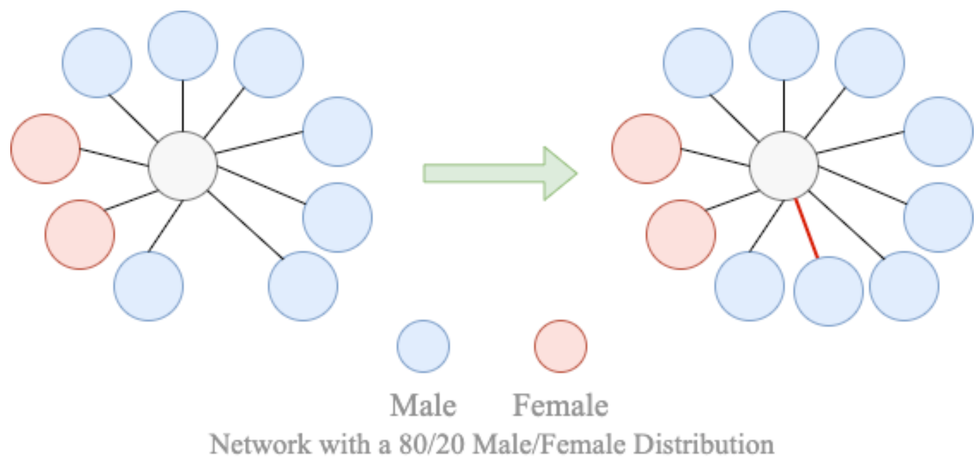
Within the *FairNet* library, we propose two independent algorithms -- *FairLabel*, and *FairEdges*.

FairLabel is intended to be used when some nodes have missing metadata. It employs a genetic algorithm to fill these nodes with the combination of labels that most reduce the number of marginalised nodes (following the above metric).

FairEdges finds a combination of edges to be added minimising the number of marginalised nodes, with relatively limited modifications to the network. Edges are selected following the ***triadic closure principle*** (e.g., we assume that a non-existing edge that would close 10 triangles is more likely to appear than one that would close 2 triangles). Plausible edges are then encoded in a binary vector. Starting from such a vector, a genetic algorithm tries to minimise the number of marginalised nodes.

Experimental results show that the *FairNet* library successfully reduces the number of discriminated nodes.

A ↔ B	C ↔ B	E ↔ H	C ↔ G	A ↔ G	A ↔ H
1	0	1	1	0	0



Network with a 80/20 Male/Female Distribution

Key bibliography
- Lerman, K., Yan, X., & Wu, Z. (2016). The "Majority Illusion" in Social Networks. PLOS ONE, 11(2), e0147617
- Sırbu, A., Pedreschi, D., Giannotti, F., et al. (2019). Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model. PLOS ONE, 14(3), e0213246.
- Karimi, F., Oliveira, M., & Strohmaier, M. (2022). Minorities in networks and algorithms. ArXiv.
- Huang, J., Gates, A. J., Sinatra, R., et al., (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. Proceedings of the National Academy of Sciences, 117(9), 4609-4616.
- Nguyen, H., Gokhale, S.S. Analyzing extremist social media content: a case study of Proud Boys. Soc. Netw. Anal. Min. 12, 115 (2022)
- Arias Ramos, N., Calvo S´anchez, M. D., Fernandez-Villa, T., et al.(2018) Social exclusion of the adolescent with overweight: study of sociocentric social networks in the classroom, Pediatric Obesity, 13, 614– 620
- Castelnovo, A., Crupi, R., Greco, G. et al. A clarification of the nuances in the fairness metrics landscape. Sci Rep 12, 4209 (2022).
- Sapiezynski, P., Stopczynski, A., Lassen, D.D. et al. Interaction data from the Copenhagen Networks Study. Sci Data 6, 315 (2019).

A project by Federico Mazzoni & Andrea Failla
federico.mazzoni1@gmail.com
andrea.failla.ak@gmail.com

Paper under review. The code will be made public upon acceptance.

Università di Pisa