The `wisdom_of_crowds`: an efficient, philosophically-validated, social epistemological network profiling toolkit

Colin Klein[(a)], Marc Cheong[(b)], Marinus Ferreira[(c)], Emily Sullivan[(d)], and Mark Alfano[(c)]

(a) Australian National University
(b) The University of Melbourne
(c) Macquarie University
(d) TU/e Eindhoven University of Technology

Background

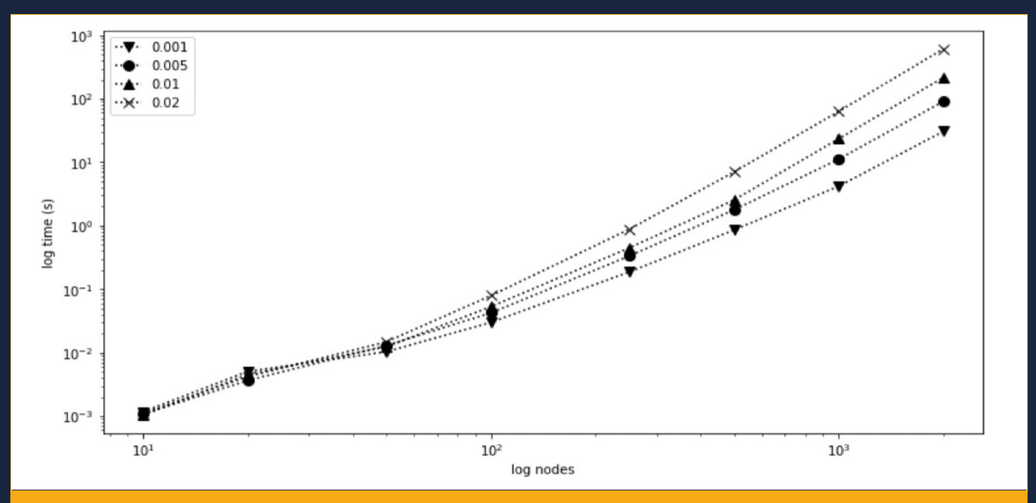
- Social epistemology: 'concerns the testimony of others embedded in social contexts' [1].
- Social epistemic networks – e.g. retweets on Twitter.
- Consider this problem: 'gossip heard from two people seems more reliable than from one, but that reliability is undermined if both heard it from the same person' [2].

State of the art

- Sullivan et al. [3] devised a method for 'quantitatively characterizing the epistemic position of individuals in a network'.
- Based on Surowiecki's [4] Wisdom of Crowds hypothesis.
- Limitation: bespoke, closed-source codebase; scaling up is needed; needs testing on robustness of code.

Our contribution:
an **open source**, **peer-reviewed**, re-implementation in Python of the core Sullivan et al. [CITE] concepts; **optimized** to deal with larger networks; built on **existing standards** to support **cross-disciplinary** collaboration.

```
pip install wisdom_of_crowds
```



Code profiling output: robust up to magnitudes of 10^3 nodes. Empirically tested on ~10^4 nodes and ~10^5 edges.

Core concepts [3]

Defining the m, k observer: we say that a node **n** is an **m, k observer** just in case it receives information from a set of at least k different nodes which are pairwise at least m steps away from one another.

S(n), independence of sources.
S(n) gives a measure of the independence of sources to node n.

$$S(n) = \begin{cases} 0 & \text{if } s = \emptyset \\ \max\{mk : (m,k) \in s\} & \text{otherwise} \end{cases}$$

D(n), diversity of sources.
D gives the number of distinct types of information that feed into n.

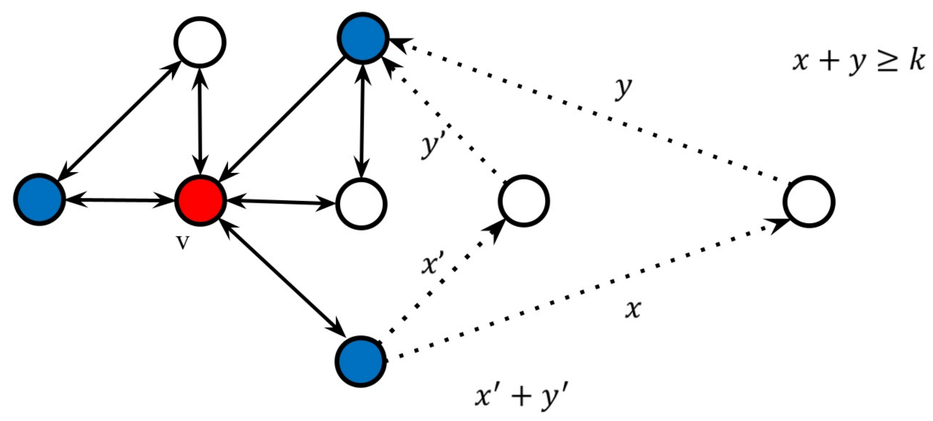
$$D(n) = \left| \bigcup \{a_i : i \in s\} \right|$$

π(n), epistemic position.
The epistemic position of a node is a function of both the diversity and independence of sources.

$$\pi(n) = S(n)D(n)$$

Worked example

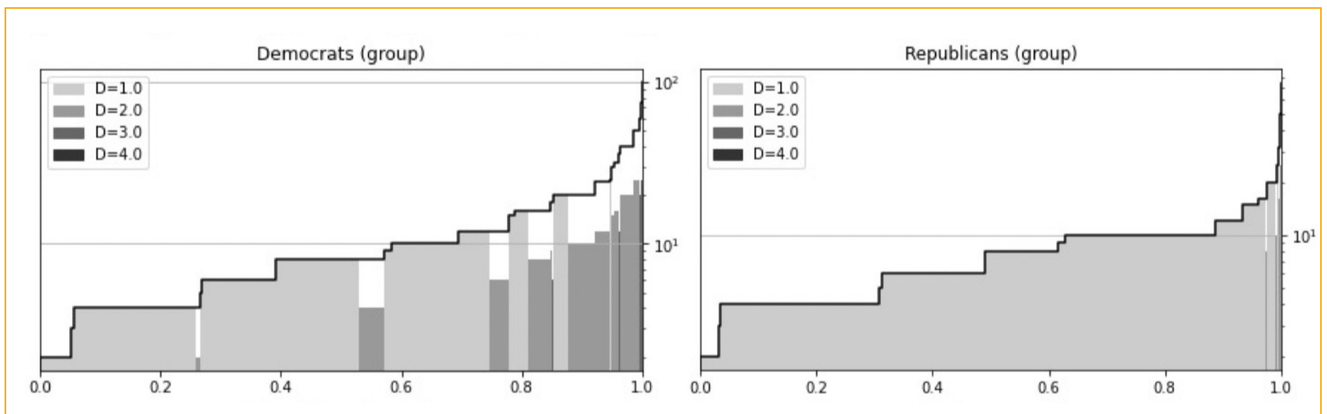
v is an **(m,k)-observer** in structure **S** if and only if
- There are **m** vertices in **S** with an edge to **v**, and
- Every **xy-path** with length < **k** contains **v**



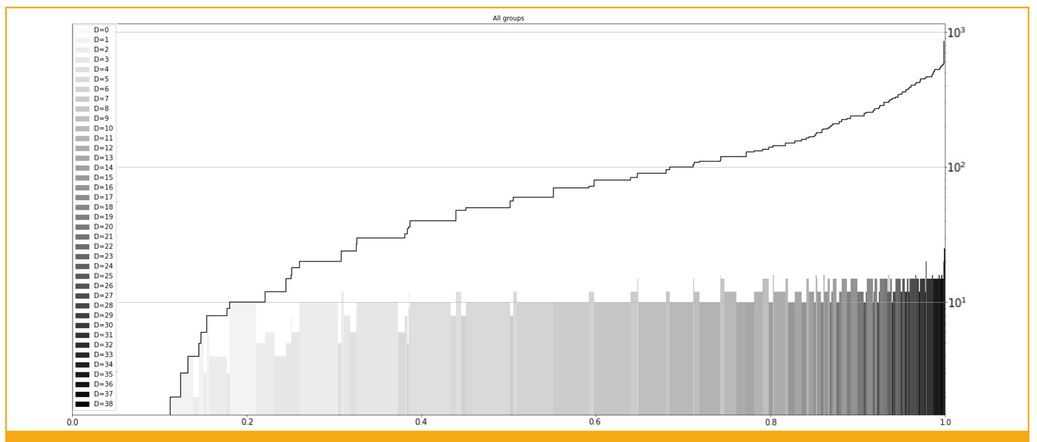
Hence, 🔴 is a (3,2)-observer.

Imagine structure S as a Twitter retweet network, with each node (@user) having a diverse set of views (e.g. 'for' and 'against')

Applications



*Information-sharing dynamics during the **Black Lives Matter movement** on **Twitter**. Republicans are a monoculture socially (low D(n)), compared to Democrats (evidence of higher D(n) values).*



***email-Eu-core** network of European researchers: researchers with contacts from a more diverse range of disciplines have higher D(n) to optimise overall **π**.*

References

1. Goldman, A. I. (1999). Knowledge in a social world. Oxford University Press
2. Alfano, M. and Robinson, B. (2017). Gossip as a burdened virtue. *Ethical Theory and Moral Practice*, 20(3):473–487
3. Sullivan, E., Sondag, M., Rutter, I., Meulemans, W., Cunningham, S., Speckmann, B., and Alfano, M. (2020). Vulnerability in social epistemic networks. *International Journal of Philosophical Studies*, 28(5):731–753.
4. Surowiecki, J. (2005). *The Wisdom of Crowds.* Abacus, London.

Marc.Cheong@unimelb.edu.au
Colin.Klein@anu.edu.au

Generously supported by:
ARC Grant DP190101507.
Templeton Grant 61378.
The Nectar Research Cloud and Melbourne Research Cloud (University of Melbourne).
[Template credits: BrightCarbon]

Scan the QR code to download the full paper/poster, or check out our Python library and source code.

linktr.ee/wisdom_of_crowds