

Excess reciprocity distorts reputation in online social networks

G. Livan, F. Caccioli, T. Aste

University College London

The digital economy is increasingly self-organizing into a "platform society" where individuals exchange knowledge, goods and resources on a peer-to-peer (P2P) basis. In recent years we have indeed witnessed how a number of well-established business-to-consumer sectors, such as the taxi and hotel industries, have been disrupted by the emergence of the novel sharing economy P2P marketplaces.

The peer-to-peer economy relies on establishing trust in distributed networked systems, where the reliability of a user is assessed through digital peer-review processes that aggregate ratings into reputation scores.

Given the expected growth of the P2P paradigm, digital reputation will increasingly become central in our online lives, as it will determine access to substantial economic opportunities. Hence, it is crucial to ensure that digital peer-review systems produce reliable reputation scores.

Being decentralised, P2P systems are often thought to promote more economic freedom and democratisation. Yet, their current lack of regulation exposes them to a number of biases which can distort their functioning. Game theoretic considerations, and plenty of anecdotal evidence, suggest that users are often incentivised to reciprocate both positive and negative ratings.

Here we present evidence of a network effect which biases digital reputation, revealing that P2P networks display exceedingly high levels of reciprocity. In fact, these are much higher than those compatible with a null assumption that preserves the empirically observed level of agreement between all pairs of nodes, and rather close to the highest levels structurally compatible with the networks' reputation landscape. This indicates that the crowdsourcing process underpinning digital reputation can be significantly distorted by the attempt of users to mutually boost reputation, or to retaliate, through the exchange of ratings. We uncover that the least active users are predominantly responsible for such reciprocity-induced bias, and that this fact can be exploited to obtain more reliable reputation estimates. Our findings are robust across different P2P platforms, including both cases where ratings are used to vote on the content produced by users and to vote on user profiles.

[1] G. Zervas, et al., working paper **SSRN**, 2366898 (2016).

[2] V. Lehdonvirta, *Policy & Internet* **3**, 263 (2015).

[3] Fagiolo et al., *J. Econ. Interact. C* **8**, 75 (2016).