A social science-grounded approach for quantifying online social influence

Rohit Ram[1] Marian-Andrei Rizoiu[1,2]
[1]The Australian National University, [2]University of Technology Sydney

Context. Social influence has become an incredibly powerful tool with the advent of social media; as such, measuring and understanding it has become an increasingly important endeavor. Understanding and measuring online social influence are essential for sociologists studying modern events, such as understanding how 'influencers' promote products effectively, how Queenslanders swung the last Australian election or how anti-vaccination sentiment spreads. Social influence has been studied thoroughly in both the computational and social science contexts, however there is often a disconnect between the two fields. Computational techniques are often not grounded in social theory, nor are they validated by humans. Conversely, social scientists rarely exploit the vast amounts of data from social media and they lack the tractable computational models to quantify influence efficiently.

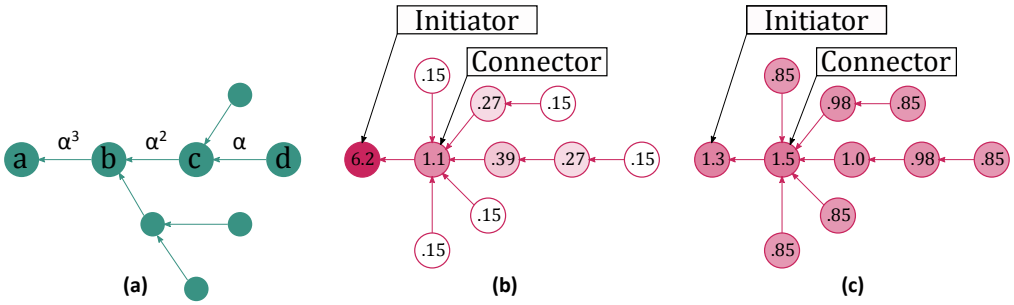

Figure 1: **Social-Temporal Decay Model.** **(a)** Depiction of the mechanism by which social credit is passed along the path *abcd* of a diffusion scenario. Edges show the retweet relations (e.g. *b* retweeted *a*. Assuming *d* has one social credit (for retweeting once), the values on the arrows show how much of it is passed up along the diffusion chain to its parents. **(b, c)** The distribution of social credit for a high value of $\alpha = 0.85$ **(b)** and a low value of $\alpha = 0.15$ **(c)**, for the same diffusion scenario. The values in the circles show the influence of each node under the two settings.

Social influence model. In this work, we propose a social science-grounded model to quantify social influence in the online setting – dubbed the *spatio-temporal model* –, constructed from the sociological properties of social influence, such as the transitivity, the asymmetry, and the composability of social influence. Starting from the observation that retweeting is considered a form of endorsement, we quantify the influence of Twitter users as the expected number of retweets they receive. First, for each retweet cascade we build the stochastic graph of retweet diffusion scenarios – i.e. the unobserved path that retweeting takes in the social network –, in which the retweet probability accounts for two known social phenomena: *preferential attachment* and the propensity of users to retweet *fresh content*. For every diffusion scenario, we compute the spatial-temporal score, which allows retweeters to pass a proportion of the social credit they earn by retweeting themselves or being retweets, $\alpha \in (0, 1)$, to their parent as shown in Figure 1. As a result, the *initiator* of the cascade receive most of the credit for higher values of α, while a lower α leads to a more evenly distributed credit among the users of the diffusion and higher scores for the *connectors* – i.e. highly followed users that generate a lot of retweets.

Matching with the human perceived social influence. To validate the model and to fit an α, we derived a score of perceived social influence for several users, via pairwise comparisons collected on a crowd-sourcing platform. Notably, we obtain a low α of 0.15, which supports the hypothesis that people perceive *connectors* as being more influential.