UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

MEASURING THE SOCIAL INFLUENCE OF ONLINE COMMUNICATIONS AT THE INDIVIDUAL AND COLLECTIVE LEVEL: A CAUSAL FRAMEWORK

by Dimitra (Mimie) Liotsiou

A central problem in the analysis of observational data is inferring and measuring causal relationships - what are the underlying causes of the observed outcomes? With the recent proliferation of Big Data from Web-mediated social communications, it has become important to measure the social influence of online communications, i.e. to determine to what extent online social communications cause certain messages, ideas, behaviours to be widely adopted (to 'go viral'), and to what extent other causes also play a role. This thesis proposes a critique and a causal conceptual and methodological framework for analysing, measuring and qualifying the social influence of online text-based communications in a given setting, while accounting for the effects of other relevant causes, at the individual and the collective level, based on 'found' observational digital data. At the individual level, this thesis demonstrates theoretically and through an analytical discussion how the proposed causal framework can successfully address the key limitations of the popular contagion-based paradigm for online social influence, enabling researchers to disentangle, measure and qualify the social influence of online communications, versus the effects of other (social and non-social) causes. At the collective level, by applying the proposed causal framework, this thesis empirically shows that the assumption of the contagion-based paradigm that the influence of online communications can be measured in isolation, without regard for other causes, does not hold, as it is empirically found that other causes can introduce non-negligible confounding bias to estimates of the social influence of online communications, and that these confounding causes themselves can be stronger causes of the outcomes of interest than online social communications, more robust to bias, with their effects following a much steadier pattern over time. Overall, the proposed causal framework enables researchers to empirically test claims and assumptions about which causes should be accounted for when measuring the social influence of online communications on outcomes of interest, and to pick apart and compare the social influence of online communications versus the influence of other causes, over time and across contexts, at the individual and at the collective level.