Network Dynamics of Knowledge Creation in the Wikishere

Myshkin Ingawale

Abstract

The rapid success of initiatives such a Source-Forge, Slashdot, Wikipedia, Digg, Facebook, Del.icio.us and others over the last decade have aroused tremendous academic interest in the field of social computing This thesis addresses the research gaps in extant literature in the understanding of growth dynamics, *emergence* and *persistence* of high quality content in a social computing forum called Wikipedia. Wikipedia is a free online encyclopedia that accepts contributions from anyone. Researchers have increasingly become interested in such platforms that focus on methods for harvesting the collective intelligence of groups of people in order to realize greater value from the interaction between users and information systems. This thesis set out with a mandate of meeting exploratory, explanatory and prescriptive research goals in its study of Wikipedia. We posit that Wikipedia is an interaction network. The rationale for this stance is grounded in the observation that in the Wiki publication model, each article is not merely a passive information object but the focal point for social interaction in the form of discussion, negotiation, conflict and collaboration between the members of the contributor community.

In our work, temporal interaction in six different independent language Wikipedias was mapped over four years: 2005, 2006, 2007 and 2008. This interaction was represented as a bipartite graph, where nodes represent users and articles and ties represent interaction ties between the two. This was then converted to two single mode interaction networks: a user network and an article network. On exploratory analysis of the network structures of these networks, we documented power law degree distribution, small-world characteristics, assortative mixing and non-hierarchical organization. We also documented a self-similar mechanism by which the giant component forms and Wikipedia's interaction graphs grow over time. Noting the importance of high degree, structural hole spanning hubs, we proposed a mechanism by which high quality emerges in Wikipedia. Furthermore, we tested and found association at the node-level between the position in the interaction network structure and the quality of output: We showed that Featured Articles occupy hub positions that span structural holes. Wikipedia is shown to be governed by social norms, that persist despite

significant churn in the contributor base. Using our results of scale-free nature and assortativity in degree distribution and drawing from prior work in diverse areas of memetics, epidemiological approach to culture, contagion theory and network science, we also put forward a network-theoretic explanation for the sustainability of governance mechanism. In particular, we claim that provided an interaction network displays the structural characteristics of scale-free degree distribution and assortative mixing of nodes, it is a candidate for persistence of social norms in its inner-core. This process is path-dependent. In a system where the initial social norms benefit quality, the governance mechanism they provide is likely to persist indefinitely. The contributions of this work are as follows:

Exploratory

- Carries out the longitudinal network analysis of 6 different language Wikipedias, and documents the evolution of structural measures over the full life-cycle, beginning from inception. Uncovers interaction networks to be scale-free and non-hierarchical.
- Develops an understanding of growth dynamics of Wikipedia and documents the self-similar nature of growth over time.
- To the best of our knowledge, this is the first work to investigate, uncover and analyze interaction network dynamics in Wikipedia or any similar social system.

Explanatory

- Uncovers the mechanism by which high quality articles emerge in an interaction network, and highlights the role played by hubs therein. Finds association between position of article in interaction network and its quality of content.
- Reconciles observations of dynamics of growth with existing theory in management science within the network paradigm, most notably small-world theories and the theory of structural holes.
- Shows Wikipedia to be governed by social norms and proposes a plausible explanation to explain the persistence of social norms in the Wikipedia community. Finds assortative mixing and scale free characteristics in Wikipedia's interaction networks

Prescriptive

- Within the context of research results obtained and underlying assumptions, prescribes guidelines for Wiki adoption and sustainability for both not-for-profits and for-profit organizations looking to harness such social computing technologies.