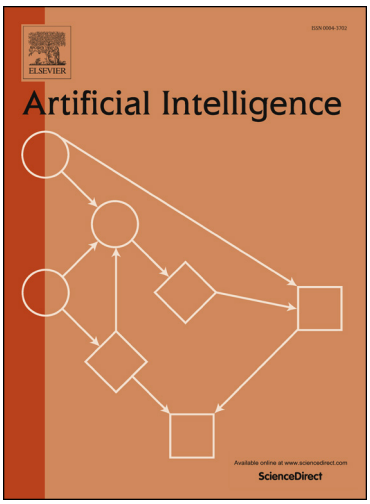
Accepted Manuscript



Diffusion centrality: A paradigm to maximize spread in social networks

Chanhyun Kang, Sarit Kraus, Cristian Molinaro, Francesca Spezzano, V.S. Subrahmanian

PII: S0004-3702(16)30071-6
DOI: http://dx.doi.org/10.1016/j.artint.2016.06.008
Reference: ARTINT 2959

To appear in: *Artificial Intelligence*

Received date: 7 February 2015
Revised date: 21 May 2016
Accepted date: 30 June 2016

Please cite this article in press as: C. Kang et al., Diffusion centrality: A paradigm to maximize spread in social networks, *Artif. Intell.* (2016), http://dx.doi.org/10.1016/j.artint.2016.06.008

Diffusion Centrality: A Paradigm to Maximize Spread in Social Networks

Chanhyun Kang[a], Sarit Kraus[b], Cristian Molinaro[c], Francesca Spezzano[d], V.S. Subrahmanian[a]

[a]*Department of Computer Science, University of Maryland, USA*
[b]*Department of Computer Science, Bar-Ilan University, Israel*
[c]*DIMES, Università della Calabria, Italy*
[d]*Department of Computer Science, Boise State University, USA*

Abstract

We propose *Diffusion Centrality* (DC) in which semantic aspects of a social network are used to characterize vertices that are influential in diffusing a property p. In contrast to classical centrality measures, diffusion centrality of vertices varies with the property p, and depends on the diffusion model describing how p spreads. We show that DC applies to most known diffusion models including tipping, cascade, and homophilic models. We present a hypergraph-based algorithm (HyperDC) with many optimizations to exactly compute DC. However, HyperDC does not scale well to huge social networks (millions of vertices, tens of millions of edges). For scaling, we develop methods to coarsen a network and propose a heuristic algorithm called "Coarsened Back and Forth" (CBAF) to compute the top-k vertices (having the highest diffusion centrality). We report on experiments comparing DC with classical centrality measures in terms of runtime and the "spread" achieved by the k most central vertices (using 7 real-world social networks and 3 different diffusion models). Our experiments show that DC produces higher quality results and is comparable to several centrality measures in terms of runtime.

Keywords:
Social networks, diffusion model, logic programming, quantitative logic

1. Introduction

An increasingly important problem in social networks (SNs) is that of assigning a "centrality" value to vertices which will reflect their importance within the SN. Well-known measures such as *degree centrality* [21, 46], *betweenness centrality* [8, 20], *PageRank* [9], *closeness centrality* [49, 5], and *eigenvector centrality* [7]