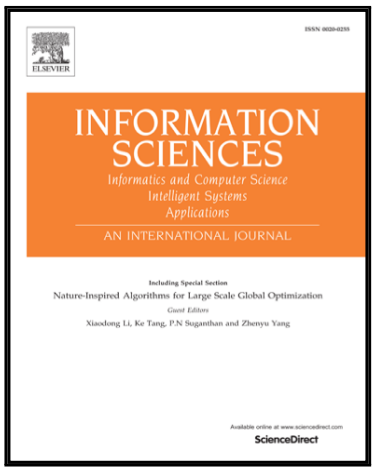

Influence Maximization in Social Networks Based on Discrete Particle Swarm Optimization

Maoguo Gong, Jianan Yan, Bo Shen, Lijia Ma, Qing Cai

PII: S0020-0255(16)30500-X
DOI: 10.1016/j.ins.2016.07.012
Reference: INS 12338

To appear in: *Information Sciences*

Received date: 10 April 2015
Revised date: 30 June 2016
Accepted date: 6 July 2016

Please cite this article as: Maoguo Gong, Jianan Yan, Bo Shen, Lijia Ma, Qing Cai, Influence Maximization in Social Networks Based on Discrete Particle Swarm Optimization, *Information Sciences* (2016), doi: 10.1016/j.ins.2016.07.012

Influence Maximization in Social Networks Based on Discrete Particle Swarm Optimization

Maoguo Gong[a,*], Jianan Yan[a], Bo Shen[a], Lijia Ma[a], Qing Cai[a]

[a] *Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China*

Abstract

Influence maximization in social networks aims to find a small group of individuals, which have maximal influence cascades. In this study, an optimization model based on a local influence criterion is established for the influence maximization problem. The local influence criterion can provide a reliable estimation for the influence propagations in independent and weighted cascade models. A discrete particle swarm optimization algorithm is then proposed to optimize the local influence criterion. The representations and update rules for the particles are redefined in the proposed algorithm. Moreover, a degree based heuristic initialization strategy and a network-specific local search strategy are introduced to speed up the convergence. Experimental results on four real-world social networks demonstrate the effectiveness and efficiency of the proposed algorithm for influence maximization.

Keywords: Social networks, Influence maximization, Cascade model, Particle swarm optimization

1. Introduction

With the great popularity of the Internet, plenty of online social networks, such as Facebook, Twitter and microblogs, have flourished in modern society. Having more than millions of users and billions of interconnections, social networks are not only serving as communication tools for users, but also potential marketing platforms for companies and advertisers [3, 13, 45]. Recent researches have shown that people place more value on recommendations from their friends than those from other channels like newspaper, TV and billboard [9, 59]. Thus, the effects of word-of-mouth, deriving from those close social circles, can make the information (i.e. interests, opinions or ideas) spread to get a wide range of cascading influence easily. Besides, compared with the traditional marketing techniques, marketing on social platforms has high profit and less investment. In recent years, many companies have chosen online social networks as marketing media to promote their new products, services or innovations. Due to limited budgets, companies are expected to select a small group of individuals as a seed set to make as many cascades as possible for the promotion [52, 57], a phenomenon called viral marketing. Influence maximization (IM) is to find a set of individuals that can generate maximum influence spread in complex networks [34, 62].

Inspired by the marketing promotion, Domingos and Richardson investigated the influence maximization as an optimization problem [15]. Kempe et al. [30, 31] pointed out that IM is essentially an NP-hard optimization problem and they proposed a greedy algorithm to solve the problem. They proved that the greedy method can guarantee $(1-1/e)$ (where e is the base of natural logarithms) optimal influence spread with respect to three commonly used information spread models, i.e., the independent cascade (IC) model [8, 20], the weighted cascade (WC) model [10, 31] and the linear threshold (LT) model [24, 48]. In order to determine an initial set with good performance in influence spread, the greedy method traverses all of nodes in the targeted social network. Besides, the computation of the influence spread for a given set is a #P-hard problem [35, 61] and the greedy algorithm combined with Monte Carlo simulations needs to run many times to obtain a precise result. The issues above result in the inefficiency of the greedy algorithm [9, 11] on the influence maximization.

*Corresponding author
Email address: gong@ieee.org (Maoguo Gong)