Near Optimal Strategies for Targeted Marketing in Social Networks

(Extended Abstract)

Rama Kumar Pasumarthi
Big Data Labs, American Express, India
ramakumar1729@gmail.com

Ramasuri Narayanam
IBM Research, India
ramasurn@in.ibm.com

Balaraman Ravindran
Indian Institute of Technology Madras
ravi@cse.iitm.ac.in

ABSTRACT

In this paper, we address the problem of Targeted Influence Maximization (TIM) through a social network. Often companies want to promote their products to certain type of customers as opposed to targeting the entire social network. That is, there is a need to maximize influence over a targeted audience in the network. Towards this end, we present a novel objective function for the targeted influence maximization problem. It turns out that this objective function is the difference between two relevant submodular functions. By building upon the recently developed theory for optimizing the difference between two submodular functions, we develop an efficient algorithm with provable guarantees. We show that the quality of solution for TIM improves using our proposed approach, when compared over a standard baseline.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*; I.2.6 [**Artificial Intelligence**]: Learning

General Terms

Machine Learning, Social Networks

Keywords

Targeted marketing, Social networks, Submodular functions, Seed selection, Greedy approximation algorithms

1. INTRODUCTION

Traditional viral marketing strategies based on influence maximization/diffusion processes focus on maximizing awareness of the product, which is equivalent to finding a set of k initial seeds to promote the product so that a large number of individuals in the network will be influenced to buy that product [2].

Most of the relevant literature on viral marketing of products attempt to maximize influence over the entire social network. However, companies often want to promote their products to certain type of customers as opposed to targeting the entire social network. For instance, the company introducing a new four-wheeler insurance plan should target only those having a four-wheeler, and not those having a two-wheeler. Therefore it is important to consider the problem of maximizing the influence over a targeted set of individuals in a social network. Towards this end, in this paper, we introduce the problem of targeted influence maximization (TIM) through a social network and present a new framework to address this problem.

Role of Incentives in Product Marketing: It is known that viral marketing works best when valuable and tangible incentives are offered. For example, an athletic clothing multichannel retailer recently offered an incentive when it launched a viral marketing campaign that rewarded message recipients with a free T-shirt and a \$1 donation to a charitable trust when an individual sent the special email message to five friends and three of those friends opted in to the retailer's catalog. It is reported that this campaign was tremendously successful leading to three times higher click-through rate than normal.

Motivated by this, we now consider a company that is interested in the targeted advertising of its product by offering incentives to those promoting the product. Since companies often have budget constraints, the company would want to minimize the incentives offered to non-targeted nodes while maximizing the influence in the targeted set A. Thus, a reasonable objective function for TIM would be to maximize the difference between the expected number of active target nodes in A and the expected number of active non-target nodes.

2. PROPOSED APPROACH TO TARGETED INFLUENCE MAXIMIZATION

Let S denote set of seed nodes. Let T correspond to the time window of observation. Let $\sigma_A(S,T)$ be the expected number of target nodes infected within time window T given that nodes in S were seed nodes. Formally,

$$\sigma_A(S,T) = E\left[\sum_{v\in A} I(t_v < T)\right] = \sum_{v\in A} Pr(t_v < T)$$

where $I(t_v < T)$ takes value 1 if node v gets activated within T, otherwise 0.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

Given a time window T, we now define the objective function for the problem of targeted influence in a given social network as:

$$\arg \max_{S:|S|\leq k} \sigma_A(S,T) - \rho \times \sigma_{V\setminus A}(S,T)$$

where ρ is a parameter that determines how strictly we want to penalize the adoption of non-target nodes. If ρ is zero, the problem reduces to the traditional influence maximization problem [2] over the target set. The higher ρ is, the more the penalty for non-target adoption.

2.1 Supermodular-Submodular procedure for optimizing the difference of submodular functions

For ease of notation, we let $f(S)$ denote $\sigma_A(S,T)$, $g(S)$ denote $\sigma_{V\setminus A}(S,T)$, and $v(S) = f(S) - g(S)$. Therefore, the objective function of TIM is equivalent to maximizing $v(.)$.

Iyer *et. al* [4] have outlined a procedure, called *Sup-Sub* procedure, for maximizing (or minimizing) the difference of two submodular functions. In this algorithm, the term $m^g_X(Y)$ represents a modular upper bound for any submodular function (g). More formally, it is given by :

$$m^g_X(Y) = g(X) - \sum_{v\in X\setminus Y} [g(X) - g(X\setminus\{v\})] + \sum_{v\in Y\setminus X} g(\{v\}) \tag{1}$$

Algorithm 1: Sup-Sub procedure

Objective: Maximise $f(X) - g(X)$
$X^0 = \phi, t = 0$
while $X^t \neq X^{t+1}$ **do**
  $X^{t+1} := \arg\max_Y f(Y) - m^g_{X^t}(Y)$
  $t = t + 1$

Note that the inner loop consists of submodular function maximization, where we can use the well known greedy algorithm, as used in [2] to optimize it.

2.1.1 Submodular maximization in the inner loop

Buchbinder *et. al.* [1] developed a randomized greedy approach for constrained submodular optimization(referred to as *RAND-GREEDY-CSO*), a $1/e$-approximation algorithm for maximizing non-monotone functions in a cardinality constrained setting, i.e, $|S| \leq K$. This can be used for the submodular maximization in the inner loop of Sup-Sub procedure (Algorithm 1).

Algorithm 2: RAND-GREEDY-CSO

Objective: Maximise $h(S)$
$S_0 = \phi, U_0 = V$
for $i = 1$ **to** k **do**
  for $u \in U_{i-1}$ **do**
    $MG(u) := h(S_{i-1} \cup \{u\}) - h(S_{i-1})$
  $M_i = \arg\max_{J\subseteq U_{i-1}, |J|=k} \sum_{u\in J} MG(u)$
  $u_i \sim Uniform(M_i)$
  $S_i = S_{i-1} \cup \{u_i\}$
  $U_i = U_{i-1} \setminus \{u_i\}$
return S_k

3. EXPERIMENTAL RESULTS

Baseline: Targeted-set restricted Discounted Maximum Degree Heuristic (TD-MDH).
We use the degree centrality of a node as a measure of it's influence. Motivated by the degree discount heuristic [3], here we present a similar approach to formulate a baseline for the TIM problem.

For each node $u \in V$, we define its degree restricted to the target set A as the number its neighbors in A. That is,

$$d^A(u) = \{v \in A \mid (u,v) \in E\} \tag{2}$$

Let S_v be the set of neighbors of node $v \in V$ in the set S. Initialising $S = \Phi$, we greedily add nodes to this set that are maximal in $d^A(v) - \big(2|S_v| + \big(d^A(v) - |S_v|\big)\,|S_v|p\big)$. p usually takes small values; $p = 0.05$ was used in the experiments.

Methodology: For each of the networks, we compare the proposed method, SupSub for TIM with the baseline TD-MDH using the value of the TIM-objective function over k initial seed nodes for $k = 5, 10, 15, 20$ and 25.

Real World Networks: We used Netscience real world network, which is a co-authorship network of scientists working on network theory and experiment, compiled by M. Newman in 2006. It has 4204 nodes and 2742 edges. The weights on the directed edges were treated as the mean of the exponential distribution for modelling the transmission function.

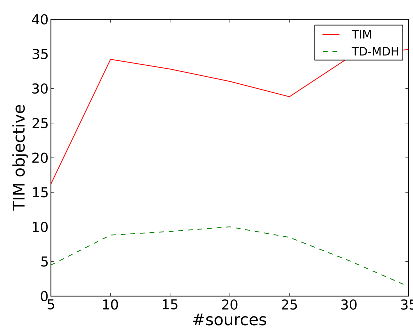

Figure 1: TIM objective on synthetic graphs for varying k using Netscience graph

REFERENCES

[1] N. Buchbinder, M. Feldman, S. Naor, and R. Schwartz. Submodular Maximization with Cardinality Constraints. In the Proceedings of 25th ACM-SIAM Symposium on Discrete Algorithms (SODA), 2014.

[2] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 137–146, 2003.

[3] W. Chen, S. Yang, and Y. Wang. Scalable Influence Maximization for Prevalent Viral Marketing in Large-Scale Social Networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2010.

[4] R. Iyer and J.A. Bilmes. Algorithms for Approximate Minimization of the Difference between Submodular Functions. In Uncertainity in Artificial Intelligence (UAI), pages 407–417, 2012.