A catchy title

Anonymous Author(s)

ABSTRACT

A convincing abstract

ACM Reference Format:

Anonymous Author(s). . A catchy title. In *Proceedings of (The Web Conference'19).* ACM, New York, NY, USA, 2 pages.

1 INTRODUCTION

- Motivations
- Background
- Contributions

2 BACKGROUND AND RELATED WORK

3 MODEL SPECIFICATION

Topic-aware influence propagation models share three main hypothesis:

i Users have different interests
ii Items have different characteristics
iii Similar items are likely to interest same users

These assumptions entirely describe a propagation cascade through a social influence which is strictly dependent on the topics and interests. To account preferences we enrich the notion of node defining it with the following triplet:

Definition 1. *(Node)* $< i,\ \mathbf{f}_i,\ \mathbf{t}_i >$

where i is the identification index while $\mathbf{f}_i$ and $\mathbf{t}_i$ are respectively the influence vector and interests vector of node i in a K dimensional topics space.

Definition 2. *(Influence vector)* $\mathbf{f}_i = [f_i^1, f_i^2, .., f_i^K]$

Definition 3. *(Interests vector)* $\mathbf{t}_i = [t_i^1, t_i^2, .., t_i^K]$

Each component represents respectively the degree of influence that the node can exerts on a particular topic and the degree of interest of the node in that topic; for these reasons these entries are non-negative defined.

The graph is defined by the set of nodes V and the set of edges $E \subseteq V \times V$: $G = (V, E)$. The topic-aware perspective implies the generalization of its adjacency matrix entry:

Definition 4. *(Link weight vector)* $\boldsymbol{p}_{uv} = [p_{vu}^1, p_{vu}^2, .., p_{vu}^K]$

where vu is the direct link from u to v, $z = 1, .., K$ is the topic, and $p_{vu}^z = t_u^z + f_v^z \in [0, +\infty)$. In this way the directed link from node v to u on a particular topic z has an importance proportional to the interest of node u on that topic, incremented by the amount of influence that node v can exert on u.

Moreover the necessity of initial conditions in the diffusion process is solved by introducing a special node in the network, called *god node*; it is connected with all other nodes and denoted with a negative index *-1*.

The Web Conference'19, May 13 - 17, San Francisco, CA, USA

.

Definition 5. *(God node)* $v_{-1} \mid E^+(v_{-1}) \equiv V \wedge E^-(v_{-1}) \equiv \emptyset$

where: $E^+(u) := \{v \in V | v \neq u \wedge e_{uv} \in E\}$ is the set of nodes followed by u, and symmetrically: $E^-(u) := \{v \in V | v \neq u \wedge e_{vu} \in E\}$.

The corpus I, i.e. the set of items with cardinality I involved in the cascades, is described by probability distributions. As a main concept LDA assumes that each topic is described by a superposition of words while a document can be statistical interpreted by a composition of topics. This particular model defines these distributions as Dirichlet distributions in order to incorporate in the model the assumption that few topics can well describe a text.

Following the standard notation of the generative LDA model we have:

- $\alpha_{z=1,..,K}$:= K-dim vector of prior topic distribution
- $\beta_{w=1,..,V}$:= V-dim vector of prior word distribution
- $\gamma_{i=1,..,|I|}$:= K-dim vector of document's distribution over topics; each entry is the probability of topic z occurring in document i: $p(k|i)$
- $\varphi_{z=1,..,K}$:= V-dim vector of topic's distribution over vocabulary words; each entry is the probability of word w occurring in topic z: $p(w|z)$

4. diffusion model: formula and definitions 5. box summary

- WoMG from thesis

4 GENERATING INTERESTS

Intro: what is the challenge here? What are we going to propose to solve it?

4.1 A propagation model

4.2 Node2vec

Simple changes:

- Prior
- Positivity

4.3 Matrix Factorization

5 IMPLEMENTATION

6 EXPERIMENTAL ASSESSMENT

RQ1: Which is the best method to represent nodes in our model? RQ2: Is this model tunable for different assumptions? RQ3: Is this model realistic?

6.1 Interest generation comparison

What is *best*?

- Scalability
- Tunable homophily

6.2 Parameters analysis

Here we fix one or two interests generation methods and we explore the range of properties for the synthetic propagation the model can generate (RQ2).

6.3 Real data

Experiments whwere we compare with real data set (RQ3).

7 CONCLUSIONS AND FUTURE WORK

REFERENCES