**Scalable Influence Maximization for prevalent viral marketing in large-scale social networks**

利用社会关系原因：

It is based on trust among individuals’ close social circle of families, friends, and co-workers. Research shows that people trust the information obtained from their close social circle far more than the information obtained from general advertisement channels such as TV, newspaper and online advertisements [16].

The problem is whom to select as the initial users so that they eventually influence the largest number of people in the network.

The main idea of our heuristic scheme is to use local arborescence2 structures of each node to approximate the influence propagation. We first compute maximum influence paths (MIP) between every pair of nodes in the network via a Dijkstra shortest-path algorithm, and ignore MIPs with probability smaller than an influence threshold θ, effectively restricting influence to a local region.

Notations

S ⊆ V be a seed set

*S*t *⊆ V* be the set of nodes that are activatedat step *t ≥ 0*

MIPG(u, v) is simply the shortest path from *u* to *v* in the weighted graph G

MIIA(MaximumInfluence In-Arborescence): the union of the maximum influence paths to *v* . It can be done using efficient implementations of Dijkstra’s shortest path algorithm.

MIOA(MaximuminfluenceOut-Arborescence): the union of the maximum influence of *v* to other nodes.When *MIIA*(*v, θ*)’s for all node *v* ∈ *V* are available, *MIOA*(*v, θ*)’s can be derived from *MIIA*(*v, θ*)’s, therefore no extra running time for *MIOA*(*v, θ*)’sis needed.

NOTE: MIIA(v, θ) and MIOA(v, θ) give the local influence regions of v, and different values of θ controls the size of these local influence regions.

pp(u, v) ：Propagation probability

σI(S) ：The **influence spread** of *S*, which is the expected number of activated nodes given seed set *S.* In[10], We define the influence of a set of nodes A, denoted *σ*(*A*), to be **the expected number of active nodes at the end of the process**, given that A is this initial active set *A*0.

*ap*(u, S, MIIA(v, θ)) : the activation probability of any node *u* in MIIA(*v, θ*)

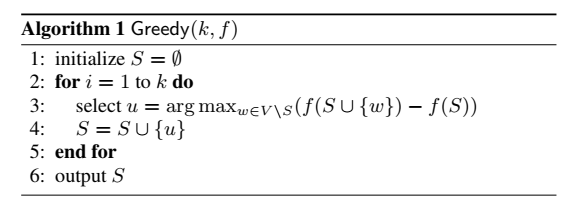
σ*M*(*S*): the influence spread of *S* in our MIA model

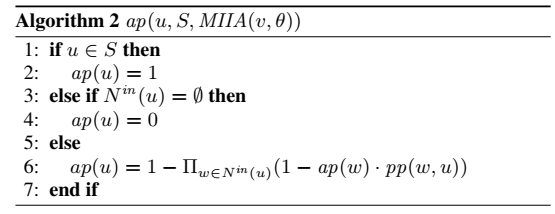
σM(*S*)= ∑v∈Vap(v, S, MIIA(v, θ))

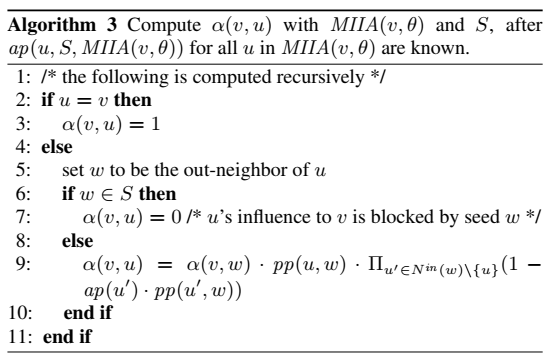
*IncInf* (*v*)：the incremental influence spread

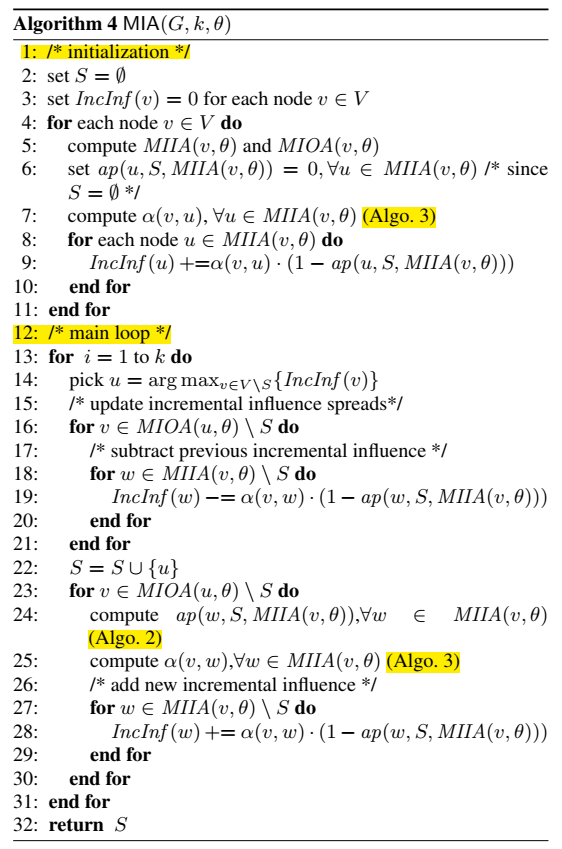
α(v, w): linear coefficient. Since ap(w) and ap(v) have a linear relationship with the linear coefficient α(v, w), the incremental influence of w on v is given by α(v, w) · (1 - ap(w)).

*For every node* v *∈* V*, our algorithm stores* MIIA*(*v, θ*),* MIOA*(*v, θ*), and for every* u *∈* MIIA*(*v, θ*), (*u, S, MIIA*(*v, θ*)) and* α*(*v, u*) are stored (note that* ap*(*u, S, MIIA*(*v, θ*)) can reuse the same entry for different seed set* S*). We also use a max-heap to store and update* IncInf (v) *for all* v *∈* V*.*









*实验：*

*对比方法*

*数据集*

*测试指标*