Appendix

Algorithm 1: Training of the fine stage

Require: number of diffusion steps T, variance schedule β_t, the whole graph $\mathcal{G}_w$

repeat
  $t \sim \text{Uniform}(1, \ldots, T)$
  calculate the time-dependent score function $\nabla_{\mathbf{i}_t} \log p(\mathbf{i}_t \mid \mathbf{i}_0)$
  Take gradient $\nabla_{\boldsymbol{\phi}} \left\| \mathbf{s}_{\boldsymbol{\phi}}(\mathbf{i}_t, t \mid \mathcal{G}^w) - \nabla_{\mathbf{i}_t} \log p(\mathbf{i}_t \mid \mathbf{i}_0) \right\|_2^2$ to update $\arg\min_{\boldsymbol{\phi}} \mathbb{E}_{p_{\text{data}}(\mathbf{i})} \mathbb{E}_{p_{\beta_t}(\mathbf{i}_t \mid \mathbf{i_0})}(\cdot)$
until converged
return $\boldsymbol{\phi}$

A Baselines

We mainly adopt the following two types of baselines to make a more comprehensive comparison. (1) Method of presupposing dissemination mode: NetSleuth makes propagation predictions under the SI model for suspected source sets to determine the most likely source set. OJC can locate possible sources under the assumption of a SIR propagation model. (2) Methods compatible with multiple modes of dissemination: LPSI proposes not to use a preset dissemination model but to use local peaks of the label propagation results as source nodes based on the proposed concept of source prominence. GCNSI uses adopted spectral domain convolution to build node representations for prediction. IVGD creates a generic framework of invertible graph dissemination models for source localization on graphs, whose major contribution is proposing a graph residual scenario to make existing graph diffusion models invertible with theoretical guarantees. SL-VAE proposed a probability model to quantify the uncertainty in the propagation process and learned a priori representation to describe the propagation pattern.

B Dataset Description

All datasets are outlined bellow:

– Digg[1] is a news aggregator, which allows people to vote on the content of web pages. For each story, it contains a list of all Digg users who have voted for the story before the data collection and the timestamp of each vote. The voter's friendship link was also retrieved.

[1] https://archive.org/details/stackexchange

- MemeTracker[2] corpus contains articles from mainstream media and blogs.
- Android[3] Android is a dataset from the Stack Exchange website where users ask and answer questions on various topics.
- Christianity[4] is also built from the Stack Exchange website, which is composed of interactions between users, and cascading represents the post sequences related to the same tag.
- Twitter[5] is a dataset of users along with tweets from their timelines with a random-walk-based crawler on the retweet graph, which is annotated as hateful or not.

Table 1: Dataset outline.

	Nodes	Edges	Cascades	Avg. cascade len
Digg	8,602	173,489	968	100.0
MemeTracker	10,244	8,417,276	15,521	15.38
Android	9,958	48,573	679	42.74
Christianity	2,897	35,624	589	27.21
Twitter	12,627	619,262	3,461	39.17

[2] http://snap.stanford.edu/memetracker
[3] https://github.com/aravindsankar28/InfVAE/tree/master/data
[4] https://github.com/aravindsankar28/InfVAE/tree/master/data
[5] https://github.com/albertyang33/FOREST/tree/master/data