

User Manual to InfluenceModels (SimPath version)

Information

- **Paper:** SimPath: An Efficient Algorithm for Influence Maximization under Linear Threshold Model. Amit Goyal, Wei Lu, Laks V.S. Lakshmanan. ICDM 2011.
- **Contact author for this package:** Wei Lu, welu@cs.ubc.ca.
- Should you have questions about this paper, you can also reach Amit Goyal at goyal@cs.ubc.ca.

Terms of Use

Redistributions and use of them in source or binary forms, with or without modification, are permitted provided that the following conditions are met:

- *For academic purposes, proper acknowledgements must be made, and the packages are used as "freeware" for which the authors do not hold any responsibility.*
- *For non-academic (commercial, or other profit-involving) purposes, a special case-by-case permission of the authors must be taken in advance.*

For the complete terms of release, please read <http://www.cs.ubc.ca/~welu/downloads.html>.

Running the Package

To compile under Linux, simply type

```
$ make
```

To run the executable, just type

```
$ ./InfluenceModels -c <config-file.txt> [...]
```

([...] means other options can be used.)

Parameters in Configuration Files (or Command Line)

In the configuration file, one can specify various parameters like input file for friendship graph, propagation log, propagation model etc. These options can also be specified on the command line. If a parameter is present in both command line and configuration file, the command line has the preference.

The Phase Parameter

One of the most important parameters is "phase". Depending on its value, different modules (algorithms) of the code are invoked.

- **SimPath** and **SPS-CELF++** both use phase 17. For **SPS-CELF++**, you shall additionally set the parameter "celfPlus" to be 1. If it's 0, **SimPath** will be executed.
- **LDAG** (by Wei Chen et al. in ICDM 2010) uses phase 15.
- Baseline algorithms, including **High Degree** and **PageRank**, use phase 20. For High Degree, set the parameter "propModel" to be HighDeg, while for PageRank, set "propModel" to PageRank.
- The basic **Greedy Algorithm with Monte Carlo** simulations, use phase 10.

Other Parameters

You will also need to use some, or all of the following parameters. Please be advised that the program might encounter segmentation faults if the needed configuration parameters are missing from either the configuration file or the command line input.

- probGraphFile : the path of the input graph file.
- budget : the number of seeds to select. The default value is 50.
- mcruns : the number of Monte Carlo simulations. The default value is 10000.
- outdir : the path of the directory storing output files. Consider to use separated folders for different kinds of experiments
- propModel : the influence propagation model that will be used by the program. IC represents Independent Cascade, LT is Linear Thresholds, and LTNew means SimPath or SPS-CELF++.
- cutoff (For SimPath and SPS-CELF++ only) : the cut-off threshold for the backtrack method in SimPath and SPS-CELF++ .
- top1 (for SimPath only) : the number of items specified for the Look Ahead Optimization in SimPath.
- tol ldag (for LDAG only): the parameter for controlling the size of LDAG. The default value is 0.003125 (1/320).