LDF: Low-Degree-Following Algorithm for Modularity Clustering

by Thang N. Dinh (tndinh@vcu.edu)

Syntax:

```
ldf   -i <input file>  [ options]
```

Find the community structure of a network specified in the input file by using Low Degree Following (LDF) algorithm. The algorithm can be found in the paper:

Thang N. Dinh, My T. Thai, Community Structure in Undirected and Directed Scale-free Networks: Approximation Algorithms for Maximizing Modularity

I.   Input format:

The <u>undirected</u> graph is given in form of adjacency list. <u>The file name's extension must be ".adj"</u>. The first line contains two numbers n and m which are the number of nodes and edges in the graph. The following m lines contain on each line a pair of integer numbers u, v specifying an edge from u to v. <u>Node indices are zero-based i.e. $0 \leq u, v \leq n - 1$.</u>

[Direction] By default, the graph is an undirected and unweighted. Loops are allowed, however, multiple edges, if specified, will be aggregated into a single edge with the combined weight. [To Implement] The program will expect a directed graph if the option **–d** is specified.

[Cost/weight]   If not otherwise specified, the graph is unweighted (or edges have uniform weight 1).   The weight/cost of edges will be expected in the input file if the option **–we** is specified.

II.   Output format:

The output file, if not specified, **is named after the input file by appending ".com".** The standard output will contain as many lines as the number of communities. Each line will contain vertices of a community, separated by one space character.

If the parameter "-clu" is specified, a cluster file in Pajek format will be produced.

III.   Parameters:
   a.   *Required parameters*

   -i *<input file>*   A file contains the adjacency list of the graph. The file must have one of the supported extension **e.g. ".adj", ".bin", etc.**

   b.   *Optional parameters:*

-o *\<output file\>*	The output file
-d	The graph is directed
-we	Supply the removal cost for each edge. Each of m lines in the input will consist of three numbers u, v, w which implies that the cost of edge (u, v) is w.
-**config** \<config file\>	Load configuration parameters from the given file. Suggested file extension: .cfg
-clu	Produce Pajek cluster file.
-v	Verbose mode, printing additional information to the standard output.

IV.   Examples:

```
ldf -i graph.adj -o graph.cs
```

Input: `graph.adj`
```
6 7
0 1
0 2
1 2
2 3
3 4
3 5
4 5
```

Output: `graph.cs`
```
0 1 2
5 4 3
```