**Overlapping Community Detection Using Seed Set Expansion**

**Version 1.1 is released on May 2, 2014.**

**The Program**

**We propose an efficient overlapping community detection algorithm using a seed set expansion approach. In particular, we develop effective seeding strategies for a personalized PageRank scheme that optimizes the conductance community score. The key idea of our algorithm is to find good seeds, and then expand these seed sets using the personalized PageRank clustering procedure.**

**Download**

**The code is released under the GNU Public License (GPL). Our seed set expansion algorithm is written in a mixture of C++ and MATLAB. The high level interface is written in MATLAB.**[**Download**](http://www.cs.utexas.edu/~joyce/codes/joyce_cikm2013/OverlapCommSSE.zip) **and extract the files. Once you prompt MATLAB, please type 'compile' inside the main directory.**

**Usage**

C = OverlapCommSSE(A, seeding, expand, k);

input:

A: adjacency matrix

seeding ('graclus\_centers' / 'spread\_hubs'): seed finding strategy

expand ('true' / 'false'): if expand is true, then the seed sets are expanded via their neighborhood

k: the number of seeds

output:

C: assignment matrix (rows represent nodes, columns represent clusters)

**Citation**

**Please acknowledge the use of the code with a citation.**

**Overlapping Community Detection Using Seed Set Expansion, Joyce Jiyoung Whang, David F. Gleich, Inderjit S. Dhillon, ACM International Conference on Information and Knowledge Management (CIKM), October 2013. [**[**pdf**](http://www.cs.utexas.edu/~joyce/paper/overlapping_commumity_cikm13.pdf)**]**

@inproceedings{jjw\_cikm13,

title ={Overlapping Community Detection Using Seed Set Expansion},

author={Joyce Jiyoung Whang and David F. Gleich and Inderjit S. Dhillon},

booktitle = {ACM International Conference on Information and Knowledge Management},

year = {2013}

}