

ECC: A Novel Clustering Algorithm for Patient Stratification

Authors: Hongfu Liu^{1,*}, Rui Zhao^{2,3,*}, Hongsheng Fang^{2,3,4}, Feixiong Cheng^{5,6}, Yun Fu^{1,7,+} & Yang-Yu Liu^{2,6,+}

¹*Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA*

²*Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA*

³*Chu Kochen Honors College, Zhejiang University, Hangzhou, Zhejiang 310058, China*

⁴*Department of Statistics, Stanford University, Stanford, CA 94305, USA*

⁵*Center for Complex Network Research and Departments of Physics, Computer Science and Biology, Northeastern University, Boston, MA 02115, USA.*

⁶*Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA*

⁷*College of Computer and Information Science, Northeastern University, Boston, MA 02115, USA*

* These two authors contributed equally to this work, + Correspondence should be directed to Y.F. (yunfu@ece.neu.edu) or Y.-Y.L. (yyl@channing.harvard.edu)

Download

[ECC_code.zip](#) (132 KB, version 1.0)

Last update: June 07, 2016

The ECC code was developed by Mr. Hongfu Liu (liu.hongf@husky.neu.edu). For any problem concerning the code, please feel free to contact Mr. Liu.

RunECC is the main function with the following inputs and outputs.

```
% Input
% IDX: the set of basic partitions
% U: the utility function (Here we use U = {'U_H', 'std', []});
% w: the weight vector of basic partitions (Default setting w = ones(r,1), r is the number of basic partitions)
% rep: the repetition time of K-means (Default setting rep = 20)
% maxIter: the maximum iteration number (Default setting maxIter = 50)
% minThres: the threshold for stopping criterion (Default setting minThres = 1e-5)
% utilFlag: the indicator to calculate the utility (Default setting utilFlag = 0)

% Output
% pi_sumbest: the objective function value of K-means
% pi_index: the partition for gene expression data
% pi_converge: the objective function value in each iteration
% pi_utility: the utility value (Here we do not calculate utility value)
% t: the execution time
```

We also provide two strategies to generate basic partitions, BasicCluster_RPS and BasicCluster_RFS. Besides, the evaluation function exMeasure including VIn, VDn, Rn and NMI, is also provided by this package

demo.m is an example.