ENSEMBLE-BASED COMMUNITY DETECTION IN MULTILAYER NETWORKS

Andrea Tagarelli, Alessia Amelio, Francesco Gullo

UNIVERSITÀ DELLA CALABRIA
Dipartimento di INGEGNERIA INFORMATICA, MODELLISTICA, ELETTRONICA E SISTEMISTICA

The 2017 European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases

ECML PKDD 2017

Experimental evaluation

Datasets

- Our experimental evaluation was mainly conducted on seven real-world multilayer network datasets

	#entities ($\|\mathcal{V}\|$)	#edges	#layers	node set coverage	edge set coverage	degree	avg. path length	clust. coeff.
AUCS	61	620	5	0.73	0.20	10.43 ± 4.91	2.43 ± 0.73	0.43 ± 0.097
DBLP	1 314 050	7 647 677	44	0.06	0.02	7.46 ± 3.06	8.59 ± 1.39	0.69 ± 0.13
EU-Air	417	3588	37	0.13	0.03	6.26 ± 2.90	2.25 ± 0.34	0.07 ± 0.08
FF-TW-YT	6407	74836	3	0.58	0.33	9.97 ± 7.27	4.18 ± 1.27	0.13 ± 0.09
Higgs-Twitter	456 631	16 070 185	4	0.67	0.25	18.28 ± 31.20	9.94 ± 9.30	0.003 ± 0.004
London	369	441	3	0.36	0.33	2.12 ± 0.16	11.89 ± 3.18	0.036 ± 0.032
VC-Graders	29	518	3	1.00	0.33	17.01 ± 6.85	1.66 ± 0.22	0.61 ± 0.89

Experimental evaluation

Datasets

- We also resorted to a synthetic multilayer network generator, **mLFR Benchmark**, mainly for our evaluation of efficiency of the M-EMCD method

- We used mLFR to create a **multilayer network with 1 million of nodes**, setting other available parameters as follows:
  - 10 layers,
  - average degree 30,
  - maximum degree 100,
  - mixing at 20% ,
  - layer mixing 2.

Experimental evaluation

Competing methods

- **flattening methods**
  - apply a community detection method on the flattened graph of the input multilayer network
  - it is a weighted multigraph having *V* as set of nodes, the set of edges, and edge weights that express the number of layers on which two nodes are connected
    - *Nerstrand* algorithm[1]

[1] D. LaSalle and G. Karypis, "Multi-threaded modularity based graph clustering using the multilevel paradigm", J. Parallel Distrib. Comput., 76:66–80, 2015.

Experimental evaluation

Competing methods

- **aggregation methods**
  - detect a community structure separately for each network layer, after that an aggregation mechanism is used to obtain the final community structure
    - *Principal Modularity Maximization* (PMM)[2]
    - *frequent pAttern mining-BAsed Community discoverer in mUltidimensional networkS* (ABACUS)[3]

[2] L. Tang, X. Wang, and H. Liu, “Uncovering groups via heterogeneous interaction analysis,” in *Proc. ICDM*, 2009, pp. 503–512.

[3] M. Berlingerio, F. Pinelli, and F. Calabrese, "ABACUS: frequent pattern mining-based community discovery in multidimensional networks", Data Min. Knowl. Discov., 27(3):294– 320, 2013.

Experimental evaluation

Competing methods

- **direct methods**
  - directly work on the multilayer graph by optimizing a multilayer quality-assessment criterion
    - *Generalized Louvain* (GL)[4]
    - *Locally Adaptive Random Transitions* (LART)[5]
    - *Multiplex-Infomap*[6]
    - *MultiGA*[7]
    - *MultiMOGA*[8]

[4] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, “Community structure in time-dependent, multiscale, and multiplex networks,” *Science*, vol. 328, no. 5980, pp. 876–878, 2010.

[5] Z. Kuncheva and G. Montana, “Community detection in multiplex networks using locally adaptive random walks,” in *Proc. ASONAM*, 2015, pp. 1308–1315.

[6] M. De Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall, "Identifying Modular Flows on Multilayer Networks Reveals Highly Overlapping Organization in Interconnected Systems", Phys. Rev. X, 5, 011027, 2015.

[7] A. Amelio and C. Pizzuti, "A Cooperative Evolutionary Approach to Learn Communities in Multilayer Networks", In Proc. PSSN, pages 222–232, 2014.

[8] A. Amelio and C. Pizzuti, "Community detection in multidimensional networks", In Proc. ICTAI, pages 352–359, 2014.

Experimental evaluation

Assessment Criteria

- Internal criteria
  - **redundancy measure**
    - actual number of redundant connections (i.e., pairs of nodes connected through edges of different layers) divided by the theoretical maximum (i.e., total number of layers times total number of node pairs in the community)
    - a global redundancy is finally obtained averaging the redundancy values over all communities
  - **multilayer Silhouette**
    - twofold modification in the definition for single-layer graphs:
      - the distance computation terms are linearly combined over all layers
      - the distance between two nodes is computed as one minus the Jaccard coefficient defined over the layer-specific sets of neighbors

Experimental evaluation

Assessment Criteria

- External criteria
  - **Normalized Mutual Information**
    - determines the alignment in terms of community memberships of nodes between a community structure and another one used as reference
    - the reference can be the solution obtained by Nerstrand on the flattened multilayer graph
    - the reference can be the layer-specific community structure solutions obtained by Nerstrand on each of the layer graphs

Experimental evaluation

Experimental settings

- The main parameter of EMCD methods, θ, was varied in its full range of admissible values, at a fine-grain step (0.001)
- We shall present results corresponding to values of θ that determined meaningful variations in terms of multilayer modularity
  - the values in the set {0.01, 0.03, 0.05, 0.07} and from 0.1 to 0.9 with step of 0.1.
- To generate the ensemble from each of the evaluation network datasets, we applied Nerstrand on the individual layer-specific graphs

Experimental evaluation

Experimental settings

- **GL** determines a community structure for each layer of a network,
  - a final solution was derived by assigning each node to the community which lays on most of the layers
- **PMM** requires an input number of communities
  - two configurations:
    1. exhaustive search for the number of communities corresponding to the best performance in terms of modularity, on every dataset
    2. input parameter set to the number of communities determined by our method
  - we set to 50 the number of runs of the k-means clustering method, whose application is required by PMM to obtain the consensus solution

Experimental settings

- **ABACUS** utilizes the *eclat* frequent-pattern mining method to generate the transactional representation of the ensemble
  - As by default configuration, the main model parameter in ABACUS (i.e., the minimum support threshold) was kept quite low on each dataset, typically in the range from three to ten

- For the genetic approaches (i.e., **MultiGA** and **MultiMOGA**), **LART**, and **Multiplex-Infomap**, we referred to the default parameters as specified in their respective works

Results

Evaluation of EMCD methods

- Modularity

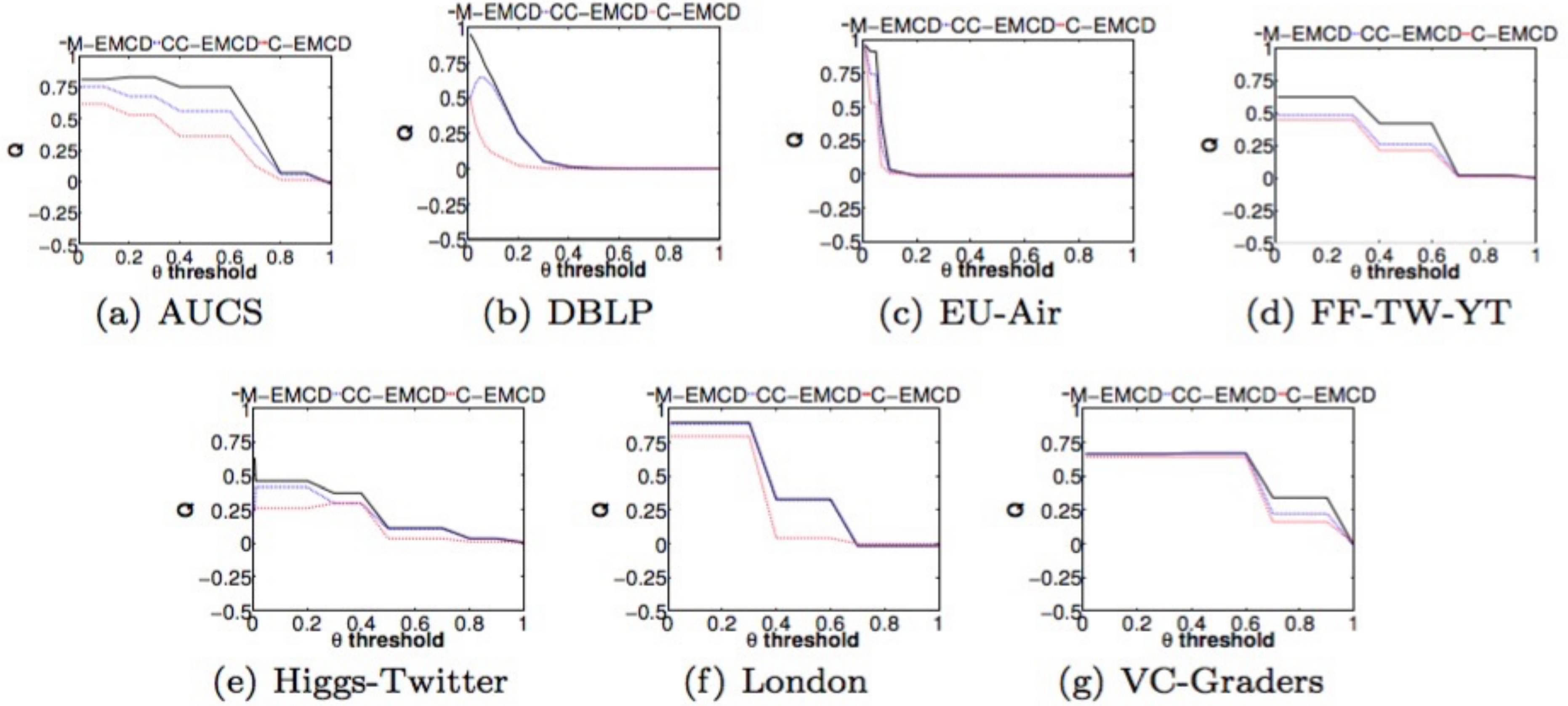


Fig. 5 Consensus solutions obtained by EMCD methods, for varying θ: modularity values. *(Best viewed in color)*

Results

Evaluation of EMCD methods

- First, the modularity value, for all methods, tends to follow a non-increasing trend as the threshold value increases
- On the contrary, the number of communities tends to increase as the threshold value becomes higher
- Among the three methods, M-EMCD turns out to be the absolute winner, reaching the highest modularity over all datasets
- Moreover, the M-EMCD solution has as good as or better modularity than that obtained by the other two methods for the same θ

Results

Evaluation of EMCD methods

Table 3 Best-modularity consensus solutions obtained by M-EMCD: modularity value with corresponding θ regime and number of communities (with percentage of singletons), and gains in modularity w.r.t. the other EMCD methods.

network	θ range	modularity	#communities (% singletons)	gain w.r.t. CC-EMCD	gain w.r.t. C-EMCD
AUCS	[0.2, 0.4)	0.863	14 (21.4%)	+0.15	+0.33
DBLP	[0.01, 0.03)	0.952	64 779 (4.0%)	+0.46	+0.46
EU-Air	[0.027, 0.07)	0.910	274 (76.6%)	+0.17	+0.38
FF-TW-YT	(0, 0.34)	0.620	86 (3.5%)	+0.14	+0.18
Higgs-Twitter	(0, 0.01)	0.625	86 (0%)	+0.39	+0.37
London	(0, 0.34)	0.895	45 (0%)	+0.01	+0.10
VC-Graders	[0.67, 1)	0.340	11 (0%)	+0.12	+0.18

Results

Evaluation of EMCD methods

- The table highlights the evident superiority of M-EMCD against the other EMCD methods
- Also, with the exception of Higgs-Twitter and DBLP, CC-EMCD tends to prevail against C-EMCD in terms of modularity
- The table also provides indications about the fraction of singleton communities in the consensus, i.e., disconnected components comprised of a single node of the graph
  - ability of M-EMCD to detect outliers in the consensus solution
- With the exception of EU-Air, the best-modularity consensus includes zero or a small fraction of singletons

Results

Evaluation of EMCD methods

- Community membership

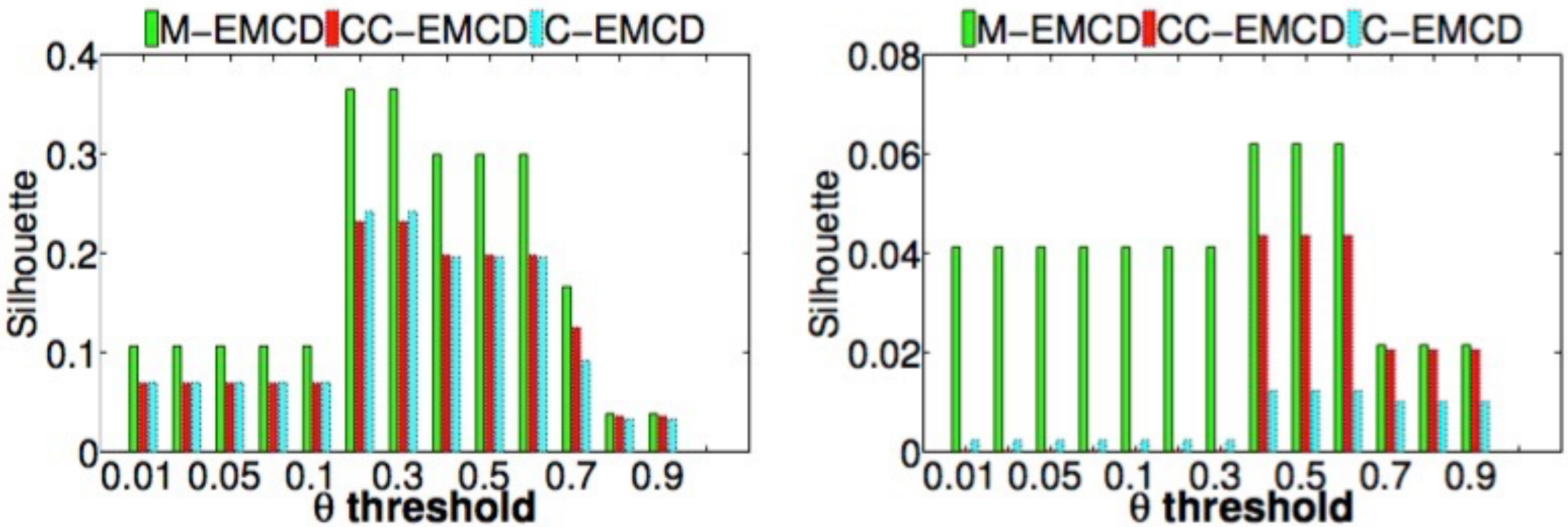


Fig. 7 Silhouette by EMCD methods for varying θ, on AUCS (left) and FF-TW-YT (right).

Results

Evaluation of EMCD methods

- Community membership

Table 4 Silhouette and global redundancy corresponding to the best consensus solutions obtained by M-EMCD, and gains w.r.t. the other EMCD methods.

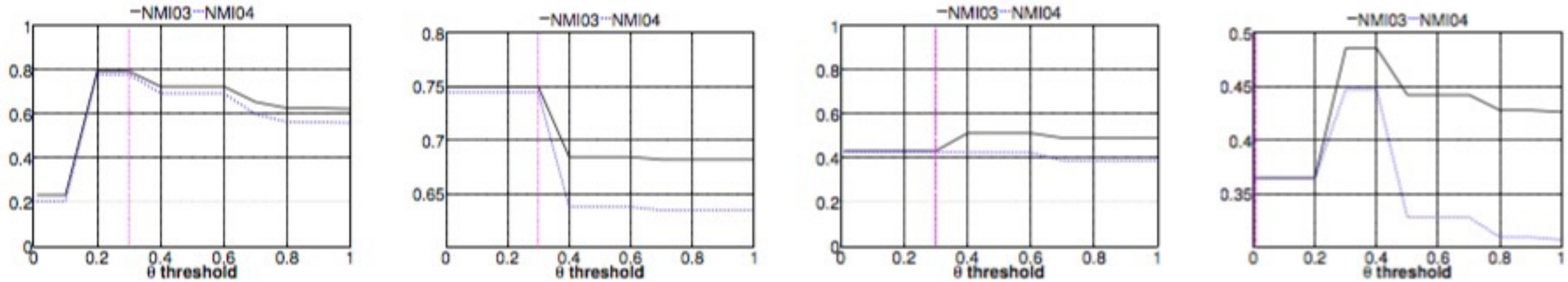
network	silhouette			redundancy		
	M-EMCD	gain w.r.t. CC-EMCD	gain w.r.t. C-EMCD	M-EMCD	gain w.r.t. CC-EMCD	gain w.r.t. C-EMCD
AUCS	0.366	+0.13	+0.81	0.910 ± 0.097	+0.04	0.0
DBLP	0.084	+0.20	+0.10	0.512 ± 0.290	0.0	-0.001
EU-Air	0.093	+0.44	+0.44	0.620 ± 0.103	+0.02	0.0
FF-TW-YT	0.041	+1.03	+0.04	0.615 ± 0.282	+0.03	0.0
Higgs-Tw.	0.052	+0.33	+0.36	0.658 ± 0.247	+0.40	0.0
London	0.179	+0.03	+0.03	0.533 ± 0.328	+0.03	0.0
VC-Graders	0.288	+0.05	+0.18	0.945 ± 0.064	+0.03	0.0

Evaluation of EMCD methods

- The silhouette of M-EMCD is higher (i.e., better) than CC-EMCD and C-EMCD over the various θ values
- In most cases M-EMCD outperforms the other methods
- Interestingly, the latter occurs consistently with the best-modularity performance
  - the largest gain in silhouette is obtained by M-EMCD over the same θ range that leads to the best modularity

Results

Evaluation of EMCD methods



(a) AUCS (b) London (c) FF-TW-YT (d) Higgs-Twitter

Fig. 8 NMI03 [29] and NMI04 [7] performances by EMCD methods w.r.t. Nerstrand on the flattened graphs, for varying θ. The vertical colored line on each plot refers to the θ value corresponding to the best-modularity consensus by M-EMCD. *(Best viewed in color)*

Table 5 Mean and standard deviation of NMI results for each dataset, between the best-modularity consensus solution by M-EMCD and the community structure by Nerstrand on each layer graph, averaged over the various layers.

	AUCS	DBLP	EU-Air	FF-TW-YT	Higgs-Twitter	London	VC-Graders
NMI03 [29]	0.752 ± 0.063	0.510 ± 0.003	0.853 ± 0.019	0.418 ± 0.040	0.305 ± 0.133	0.779 ± 0.013	0.707 ± 0.011
NMI04 [7]	0.741 ± 0.059	0.481 ± 0.003	0.847 ± 0.019	0.358 ± 0.050	0.259 ± 0.123	0.766 ± 0.024	0.667 ± 0.014

Results

Evaluation of EMCD methods

- The two NMI measures behave similarly, possibly by a scaling factor, on most θ regimes
- The highest NMI values do not necessarily correspond to the θ value by which the best-modularity consensus was obtained
- It indicates that the community membership in the solution by Nerstrand on the flattened graph can be quite different from that in the modularity-based optimal structure of consensus obtained by M-EMCD
- Also, the community membership of nodes in the consensus keeps a moderate similarity with the community memberships over each layer on average

Results

Evaluation of EMCD methods

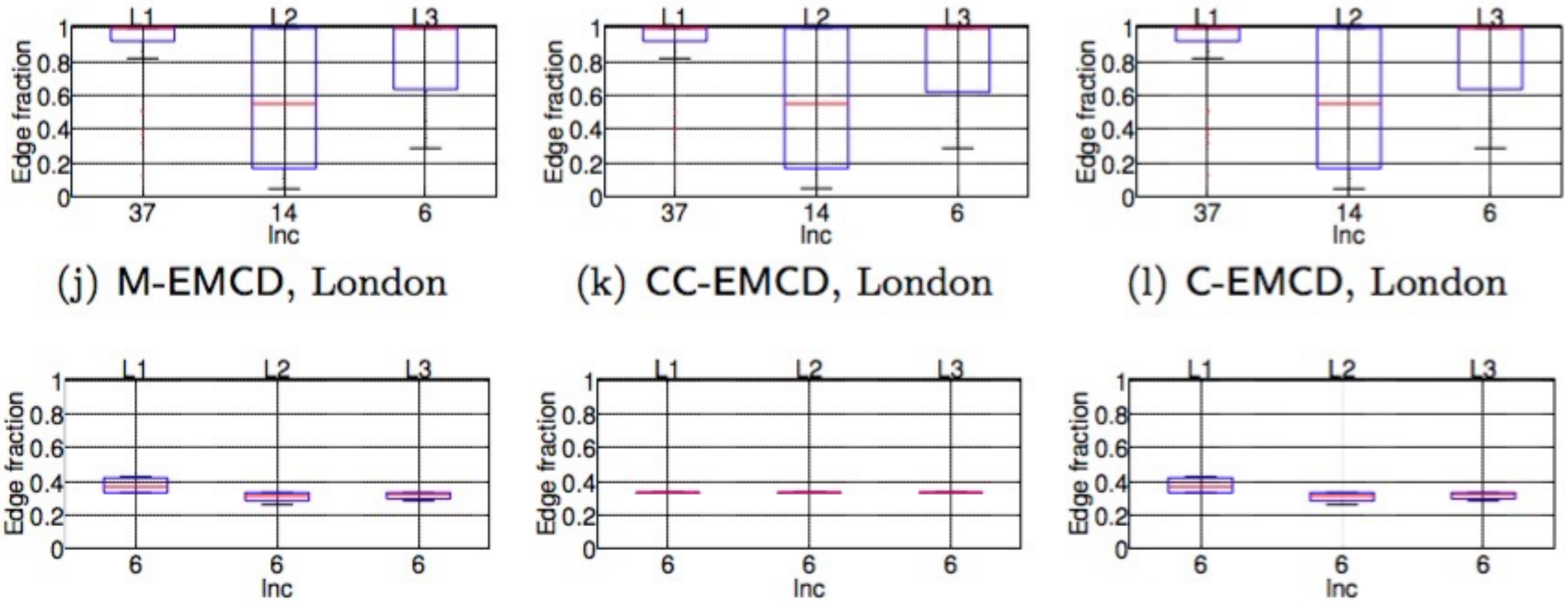
Layer coverage

- M-EMCD is able to produce consensus communities whose internal connectivity is, on average, characterized by most of the layers

- M-EMCD has also the same ability in terms of redundancy as C-EMCD, whose solution indeed represents the topological upper bound, for a given θ, of the communities being identified

Results

Evaluation of EMCD methods



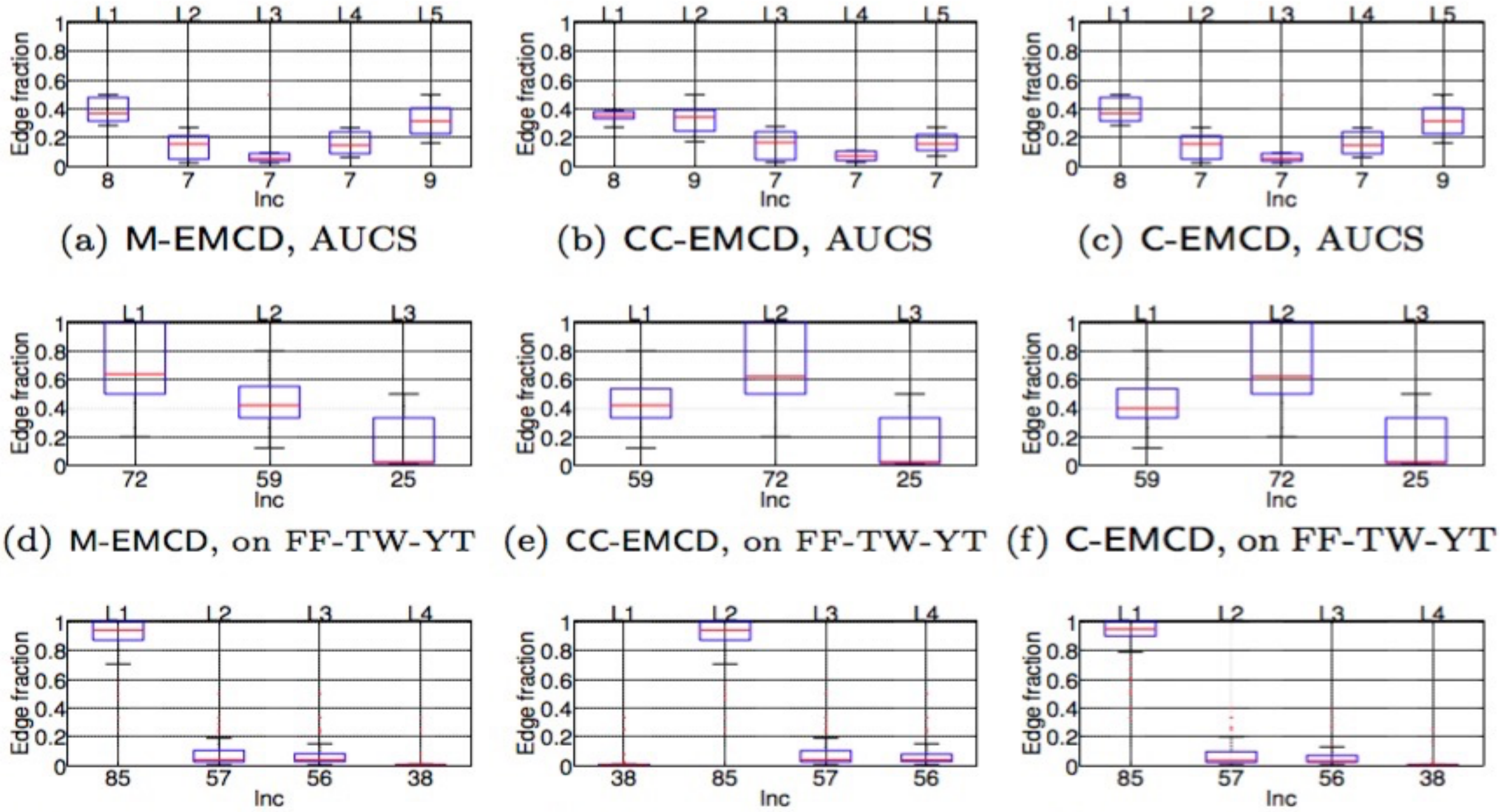
(j) M-EMCD, London (k) CC-EMCD, London (l) C-EMCD, London

(m) M-EMCD, VC-Graders (n) CC-EMCD, VC-Graders (o) C-EMCD, VC-Graders

Fig. 9 Per-layer distribution of edges over the consensus communities obtained by EMCD methods. Each EMCD solution is taken at the θ value for which M-EMCD reaches the maximum modularity. The bottom x-axis indicates, for each layer, the number of communities which contain only edges from that layer. *(Best viewed in color)*

Results

Evaluation of EMCD methods



(a) M-EMCD, AUCS (b) CC-EMCD, AUCS (c) C-EMCD, AUCS

(d) M-EMCD, on FF-TW-YT (e) CC-EMCD, on FF-TW-YT (f) C-EMCD, on FF-TW-YT

(g) M-EMCD, Higgs-Tw. (h) CC-EMCD, Higgs-Tw. (i) C-EMCD, Higgs-Tw.

Results

Evaluation of EMCD methods

- The per-layer boxplots for M-EMCD are quite similar to those for C-EMCD
- Coupling redundancy results from Table 4 and results shown in this figure, it should be noted that the highest values of redundancy of M-EMCD, observed in AUCS (0.91) and VC-Graders (0.95), correspond to situations in which the distribution of layer-characteristic communities is more uniform

Results

Evaluation of EMCD methods

- On Higgs-Twitter, there is one layer predominant on the others
- Conversely, on DBLP, all layers participated almost equally in the edge distribution of the consensus communities
- On London, the mid value of redundancy (0.533) should be reconsidered as actually all three layers participate well in the composition of the communities (the first and third layers are highly characteristic for all communities, and the second one corresponds to a distribution with median of 0.6; cf. Fig. 9-j)

Results

Evaluation of EMCD methods

- **Robustness against ensemble perturbations**

  - We configured it by specifying the number of desired communities as input parameter, rather than leaving Nerstrand free to automatically determine the number of communities

  - For a given dataset network, we generated multiple (e.g., 50) ensembles, by varying each time the setting of the number of communities to obtain on each layer of the network

    - if we indicate with $k_1,\ldots,k_l$ the number of communities Nerstrand would automatically detect, we selected the number of communities to obtain at the *i*-th layer graph (i = 1..l) by picking it in the interval $[k_i - \varepsilon, k_i + \varepsilon]$ uniformly at random, where ε is an offset selected empirically

Results

Evaluation of EMCD methods

- We report results on EU-Air since it has much more layers than the other datasets but DBLP, however unlike the latter, there is no excessive proliferation in the number of consensus communities
- We carried out 50 runs and analyzed the distribution of performance scores corresponding to the 50 ensembles
  - We perturbed the size of each layer in the ensemble at 5% of the size of the consensus solution obtained by M-EMCD (with the default configuration of Nerstrand), i.e., we set $\varepsilon = 0.05 \times |C_*| \approx 15$
- Results revealed a good robustness of M-EMCD to variations in the size of the ensemble clusterings

Results

Evaluation of EMCD methods

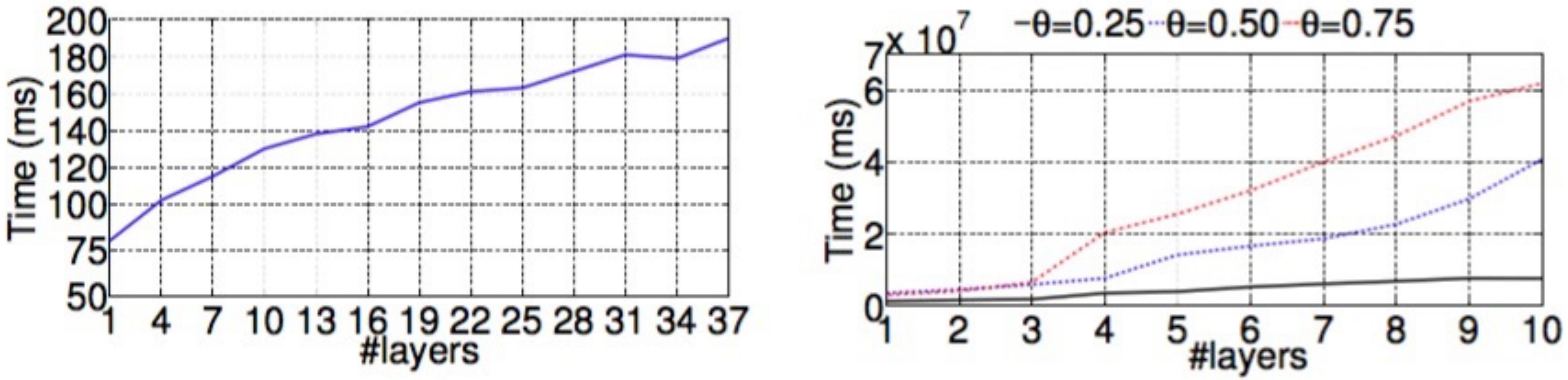
- **Efficiency evaluation**

- We focused our evaluation on two networks: EU-Air and mLFR-1M

- For each of the two network datasets, we ordered the layer graphs by increasing size, then we derived several subsets by grouping the layer graphs according to their size order
- For every subset considered, the ensemble corresponded to the community structures of the layer graphs belonging to the subset

Results

Evaluation of EMCD methods



(a) (b)

Fig. 10 Time performance of M-EMCD on (a) EU-Air and (b) mLFR-1M.

Results

Evaluation of EMCD methods

- The time performance trend grows linearly with the size (in terms of layers, hence edge set) of the network under consideration
- Therefore, our M-EMCD method scales well by increasing the size of the network
- Note also that in Fig. 10(b) the slope of the trend line tends to increase with θ, which might imply an increase in the number of consensus communities
- It should also be noted that the number of iterations, required by M-EMCD to converge, turns out to be small

Results

Comparison with competing methods

method	criterion	AUCS	DBLP	EU-Air	FF-TW-YT	Higgs-Tw.	London	VC-Graders
Nerstrand	modularity	+0.34	+0.17	+0.62	+0.24	+0.02	+0.07	+0.17
	silhouette	+0.15	+0.001	+0.01	+0.02	-0.02	+0.11	+0.01
	redundancy	+0.11	-0.12	+0.29	-0.09	-0.36	+0.17	+0.02
	#communities	+9	+13 466	+268	+43	+63	+29	+9
ABACUS	modularity	+0.10	*na*	+0.02	+0.16	+0.32	+0.10	-0.30
	silhouette	+0.38	*na*	+0.12	+0.04	+0.20	+0.24	-0.71
	redundancy	+0.20	*na*	+0.27	+0.13	+0.95	+0.39	+0.12
	#communities	+12	*na*	+250	+84	+36	+29	+10
PMM^{k^*}	modularity	+0.67	*na*	+0.89	+0.52	+0.60	+0.69	+0.24
	silhouette	+0.22	*na*	+0.23	+0.05	-0.02	+0.11	+0.13
	redundancy	-0.003	*na*	-0.07	+0.04	-0.36	-0.18	+0.003
PMM	modularity	+0.15	*na*	+0.54	+0.39	+0.26	+0.27	+0.16
	silhouette	+0.37	*na*	+0.25	+0.69	+0.20	+0.43	+0.24
	redundancy	+0.14	*na*	+0.06	+0.10	+0.79	-0.19	+0.06
	#communities	+12	*na*	+269	+76	+76	+5	+9
GL	modularity	+0.21	*na*	+0.46	+0.20	*na*	+0.62	+0.13
	silhouette	+0.17	*na*	+0.04	+0.11	*na*	+0.14	+0.10
	redundancy	+0.11	*na*	+0.32	-0.12	*na*	-0.03	+0.07
	#communities	+8	*na*	+262	-626	*na*	-212	+9

Results

Comparison with competing methods

method	criterion	AUCS	DBLP	EU-Air	FF-TW-YT	Higgs-Tw.	London	VC-Graders
Infomap	modularity	+0.50	*na*	+0.30	+0.29	*na*	+0.45	+0.26
	silhouette	+0.53	*na*	+0.20	+0.88	*na*	+0.33	+0.45
	redundancy	+0.15	*na*	+0.06	-0.33	*na*	-0.48	+0.00
	#communities	+3	*na*	+272	-117	*na*	+43	+1
LART	modularity	+0.58	*na*	+0.91	*na*	*na*	+0.89	+0.12
	silhouette	+0.50	*na*	+0.14	*na*	*na*	+0.18	+0.32
	redundancy	+0.13	*na*	+0.37	*na*	*na*	+0.53	+0.06
	#communities	-13	*na*	-107	*na*	*na*	-294	+5
MultiGA	modularity	+0.17	*na*	+0.25	*na*	*na*	+0.10	+0.16
	silhouette	+0.37	*na*	+0.06	*na*	*na*	+0.23	+0.24
	redundancy	+0.10	*na*	+0.34	*na*	*na*	-0.07	+0.06
	#communities	+9	*na*	+269	*na*	*na*	+16	+9
MultiMOGA	modularity	+0.29	*na*	+0.27	+0.40	*na*	+0.39	+0.00
	silhouette	+0.34	*na*	+0.14	+0.74	*na*	+0.21	+0.43
	redundancy	+0.08	*na*	+0.35	-0.03	*na*	+0.04	+0.01
	#communities	+7	*na*	+269	-129	*na*	+32	+7

Results

Comparison with competing methods

- Looking at modularity results, M-EMCD outperformed all competing methods
- Also in terms of silhouette, M-EMCD tends to outperform all competing methods
- Considering global redundancy values, M-EMCD generally shows higher values than those of competitors over the various networks

Results

Comparison with competing methods

- M-EMCD obtains higher global redundancy w.r.t. ABACUS and LART, and lower redundancy than communities produced by the other methods

- Coupled with modularity and silhouette results, this suggests that M-EMCD can utilize less information from the various layers than other methods to obtain higher quality consensus community structures

- M-EMCD produces much more communities than Nerstrand, ABACUS, PMM, MultiGA and MultiMOGA, while different relative behaviors correspond to comparison with the other methods on some networks

Results

Comparison with competing methods

- All methods but Nerstrand incurred memory issues on some datasets
- Some competitors methods inherently suffer from efficiency and scalability issues
  - the two genetic methods MultiGA and MultiMOGA have high computational complexity
  - LART requires the computation of similarity matrix from the pair-wise transition probabilities, and hence could not scale well with large multilayer networks
- By comparing the runtimes obtained by the competing methods with those obtained by M-EMCD, we found that M-EMCD outperforms the competing methods in terms of efficiency as well

Summary of findings

- The modularity-based approach to the EMCD problem is highly effective in producing consensus communities with improved modularity w.r.t. the CC-EMCD and C-EMCD methods

- M-EMCD also outperforms CC-EMCD and C-EMCD in terms of silhouette of community membership

- Internal connectivity of the M-EMCD consensus communities is characterized by the presence of most of the layers
  - M-EMCD has the same ability in terms of redundancy as C-EMCD

Summary of findings

- M-EMCD is relatively robust to the presence of disconnected components in a multilayer graph, as its solutions tend to have a small number of singleton communities

- Our method is relatively robust against perturbations in the input ensemble, in terms of size of its constituting clusterings

- M-EMCD scales well with the size of a multilayer network, in accordance to its computational cost that is linear in the number of edges

Results

Summary of findings

- M-EMCD consensus communities have shown to be substantially better than those generated by the competing methods, in terms of both modularity and silhouette of community membership

- Also, the method tends to use less information from the layers of the network than the competing methods, while producing better consensus community structures