On Regularizing Multiple Clusterings for Ensemble Clustering by Graph Tensor Learning: Supplementary Material

Man-Sheng Chen
Sun Yat-sen University
Guangzhou, China
chenmsh27@mail2.sysu.edu.cn

Jia-Qi Lin
Sun Yat-sen University
Zhuhai, China
linjq56@mail2.sysu.edu.cn

Chang-Dong Wang*
Sun Yat-sen University
Guangzhou, China
changdongwang@hotmail.com

Wu-Dong Xi
NetEase
Guangzhou, China
xiwudong@corp.netease.com

Dong Huang
South China Agricultural University
Guangzhou, China
huangdonghere@gmail.com

1 ADDITIONAL EXPERIMENTAL RESULTS

This supplementary material aims to provide additional experimental results for the paper titled On Regularizing Multiple Clusterings for Ensemble Clustering by Graph Tensor Learning (GTLEC). In the following, additional experimental results and analyses are provided.

Comparison Experiments *w.r.t.* ARI. The clustering results *w.r.t.* ARI on the seven datasets are shown in Table 1. From this table, we can also observe that the proposed GTLEC method can achieve almost the best clustering performance on the benchmark datasets, demonstrating the effectiveness of GTLEC from diverse evaluation metrics.

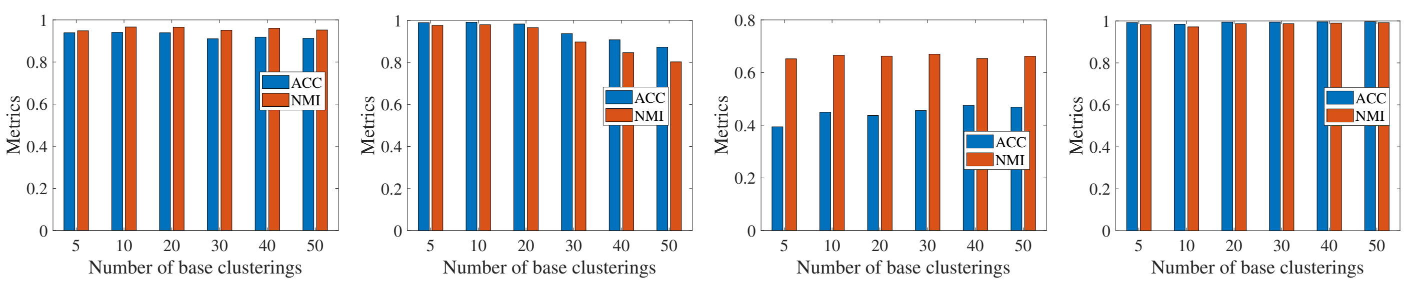
Performance with Different Number of Base Clusterings. The influence about different number of base clusterings is studied by setting its value range in $\{5, 10, 20, 30, 40, 50\}$ on the first four datasets. Specifically, the corresponding results *w.r.t.* ACC and NMI are illustrated in Figure 1. From the figure, it can be observed that the performance curves stay relatively stable when the number of base clusterings is increasing, and sometimes larger number of that may not produce better performance.

Running Time. Comprehensively, to show the time costs of different methods, the comparison results on the seven real-world datasets in terms of running time in seconds are provided in Table 2. In this paper, we mainly focus on the robustness and effectiveness of ensemble clustering. It is a promising direction to extend the current framework to large-scale applications in the future work.

*Corresponding author

Table 1: Performance comparison *w.r.t.* ARI between our method and eight ensemble clustering methods in literature.

Datasets	T-AL	T-CL	PTGP	LWEA	LWGP	E-HC	E-MC	DREC	SPCE	SEC	TRCE	CELTA	ECMS	GTLEC
Binalpha	0.3014	0.2977	0.3000	0.2788	0.2900	0.2723	0.2988	0.3060	0.2278	0.2521	0.2895	0.8343	0.2892	**0.9089**
Semeion	0.4450	0.4497	0.4451	0.4637	0.4536	0.3512	0.4368	0.4374	0.3808	0.3629	0.4507	0.9704	0.4623	**0.9828**
Caltech20	0.3695	0.3378	0.3323	**0.4247**	0.3377	0.3675	0.2567	0.2812	0.2061	0.2184	0.2344	0.3133	0.5261	0.3363
MF	0.4983	0.5019	0.5042	0.5026	0.5313	0.4163	0.4646	0.4634	0.4145	0.3577	0.4570	0.9366	0.4852	**0.9900**
MNIST	0.4609	0.4687	0.4627	0.4644	0.4479	0.4429	0.4314	0.4647	0.3520	0.2934	0.4220	0.9278	0.4885	**0.9762**
Texture	0.4962	0.4998	0.5102	0.6246	0.5789	0.5008	0.5472	0.5202	0.4242	0.3958	0.5204	0.9544	0.5556	**0.9987**
ISOLET	0.4935	0.4856	0.4953	0.5083	0.4898	0.4044	0.5015	0.4894	0.4141	0.4163	0.4973	0.5947	0.4771	**0.6679**



(a) Binalpha (b) Semeion (c) Caltech20 (d) MF

Figure 1: Performance *w.r.t.* ACC and NMI with different number of base clusterings on the first four datasets.

Table 2: Comparison results on seven real-world datasets in terms of running time in seconds.

Datasets	Binalpha	Semeion	Caltech20	MF	MNIST	Texture	ISOLET
T-AL	3.20E-03	1.88E-02	2.80E-03	1.80E-03	6.10E-03	3.30E-03	2.09E-02
T-CL	6.90E-03	1.89E-02	2.70E-03	1.40E-03	3.90E-03	5.20E-03	2.08E-02
PTGP	3.80E-02	7.04E-02	2.80E-02	2.08E-02	4.40E-02	3.81E-02	9.62E-02
LWEA	0.130E-02	0.160E-02	0.320E-02	0.290E-02	0.320E-02	0.670E-02	0.510E-02
LWGP	0.444E-01	0.993E-01	0.438E-01	0.307E-01	0.566E-01	0.519E-01	0.647E-01
E-HC	0.30E+00	0.174E+00	0.45E+00	0.35E+00	1.47E+00	1.66E+00	2.65E+00
E-MC	0.275E+00	0.074E+00	0.225E+00	0.179E+00	0.306E+00	0.432E+00	0.159E+00
DREC	1.86E+00	0.688E+00	1.21E+00	0.469E+00	3.03E+00	1.17E+00	2.78E+00
SPCE	8.22E+00	8.46E+00	3.33E+01	1.90E+01	2.60E+02	3.76E+02	8.90E+02
SEC	0.350E-01	0.513E+00	0.394E-01	0.282E-01	0.614E-01	0.600E-01	0.130E+00
TRCE	8.19E+00	6.83E+00	2.80E+01	2.41E+01	2.79E+02	2.26E+02	4.24E+02
CELTA	5.75E+01	7.22E+01	1.41E+02	1.22E+02	6.61E+02	5.15E+02	1.69E+03
ECMS	5.46E+00	1.19E+01	2.27E+01	1.88E+01	1.75E+02	2.21E+02	5.56E+02
GTLEC	6.82E+01	1.06E+02	2.73E+02	1.66E+02	3.77E+03	4.81E+03	1.36E+04