SEMI-SUPERVISED SPECTRAL CLUSTERING USING THE SIGNED LAPLACIAN

Thomas Dittrich, Peter Berger, and Gerald Matz

Institute of Telecommunications
Technische Universität Wien, (Vienna, Austria)
Email: firstname.lastname@nt.tuwien.ac.at

ABSTRACT

Data clustering is an important step in numerous real-world problems. The goal is to separate the data into disjoint subgroups (clusters) according to some similarity metric. We consider spectral clustering (SC), where a graph captures the relation between the individual data points and the clusters are obtained from the spectrum of the associated graph Laplacian. We propose a semi-supervised SC scheme that exploits partial knowledge of the true cluster labels. These labels are used to create a modified graph with attractive intra-cluster edges (positive weights) and repulsive inter-cluster edges (negative weights). We then perform spectral clustering using the signed Laplacian matrix of the resulting signed graph. Numerical experiments illustrate the performance improvements achievable with our method.

1. INTRODUCTION

Background. In many practical applications, data is gathered that can naturally be partitioned into disjoint groups such that data points within a group are similar (in a particular application-specific sense) whereas data points in distinct groups are dissimilar. These groups are referred to as clusters. Determining these clusters is thus an important problem (see e.g. [1]). Since the amount of data is often very large, an efficient data representation is desirable. We consider problems in which the (dis)similarity of data items is captured in terms of the (usually sparse) weighted adjacency matrix of a graph. Graph learning and clustering techniques have become widely popular [2–5].

A well-known graph approach for identifying clusters is spectral clustering (SC) [6–8], where the spectrum of the graph Laplacian is used to determine clusters of nodes. In some applications, the graph is not given directly but needs to be learned from data [2, 5, 9–13]. SC can be viewed as a relaxation of the computationally much harder combinatorial problem of finding the minimum ratio cut (RC) of a graph. While initial work was aimed at the case of two clusters, [14] considered an extension of SC for multiple clusters.

In some applications, cluster labels are known for a subset of the data points/graph nodes (think of a graph constructed from friendship relations in an online social network where political preference is known for a few users). In such a semi-supervised setting, the known labels can be incorporated in SC via suitable modifications of the graph's edge weights. Specifically, [15] proposed an extension of SC where the edge weight between two nodes known to belong to the same cluster is set to 1 whereas for nodes known to belong to distinct clusters the edge weight is set to 0. A method for constrained SC where "must link" and "cannot link" information in form of an additional constraint matrix is used to improve the clustering solution was introduced in [16]. In graph signal processing parlance [17, 18], such semi-supervised clustering tasks can be viewed as reconstruction problems for binary graph signals (the cluster labels) from signal samples taken on a small set of nodes.

Contributions. In this paper we propose an extension of [15] that uses negative weights for dissimilar data known to lie in distinct clusters. The negative weights lead to a signed graph and necessitate the use of the signed Laplacian matrix to perform SC, thereby promoting dissimilar nodes to end up in distinct clusters. This approach is shown to be a surrogate for the minimization of the signed RC. We discuss the specific choice of the edge weights for the known cluster labels and we demonstrate that our method achieves superior performance at the same computational complexity as competing schemes. Throughout the paper, we restrict to the case of two clusters.

2. SPECTRAL CLUSTERING REVISITED

We first give a short review of SC as background for our proposed modifications. Further details can be found, e.g., in [7, 8] (for unsigned graphs) and in [19] (for signed graphs).

2.1. Spectral Clustering on Unsigned Graphs

Consider an unsigned graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ with vertex set $\mathcal{V} = \{1, \ldots, N\}$, edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, and edge weight matrix $\mathbf{W}$, whose elements $W_{ij} \geq 0$ describe the strength of the link between nodes i and j. In many cases, this graph has been learned from the actual data [5]. Our goal is to find

Funding by WWTF Grant ICT15-119.

the two clusters, i.e., to meaningfully split the node set into two disjoint nonempty subsets $\mathcal{V}_1$, $\mathcal{V}_2$ (i.e., $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ and $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$). Equivalently, we want to determine a label vector $\mathbf{l} = (l_1, \ldots, l_N)$ such that $l_i = k$ for $i \in \mathcal{V}_k$, $k = 1, 2$. One reasonable approach attempts to determine the clusters such that the RC

$$\rho(\mathcal{V}_1, \mathcal{V}_2) = \frac{\gamma(\mathcal{V}_1, \mathcal{V}_2)}{|\mathcal{V}_1|\,|\mathcal{V}_2|} \tag{1}$$

is minimized. Here, $\gamma(\mathcal{V}_1, \mathcal{V}_2)$ is the weight of the cut-set,

$$\gamma(\mathcal{V}_1, \mathcal{V}_2) = \sum_{i \in \mathcal{V}_1} \sum_{j \in \mathcal{V}_2} W_{ij}, \tag{2}$$

and $|\mathcal{V}_k|$ is the cardinality of $\mathcal{V}_k$. Minimization of the RC favors cuts of edges with small weights (numerator) and promotes balanced cuts with clusters of similarly size (denominator).

Minimization of $\rho(\mathcal{V}_1, \mathcal{V}_2)$ is a (generally NP-hard) combinatorial optimization problem. The RC can be reformulated in terms of the graph Laplacian

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \tag{3}$$

with the diagonal degree matrix

$$\mathbf{D} = \operatorname{diag}\{d_1, \ldots, d_N\}, \qquad d_i = \sum_{j=1} W_{ij}. \tag{4}$$

Let us define the binary graph signal $\mathbf{x} = (x_1, \ldots, x_N)$ with elements

$$x_i = \begin{cases} a, & i \in \mathcal{V}_1, \\ -b, & i \in \mathcal{V}_2, \end{cases} \tag{5}$$

where a and b are suitably chosen constants. It can then be shown that

$$\rho(\mathcal{V}_1, \mathcal{V}_2) \propto \frac{1}{2\|\mathbf{x}\|^2} \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}(x_i - x_j)^2 = \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \tag{6}$$

Minimizing the RC is therefore equivalent to minimizing the Laplacian form $\mathbf{x}^T\mathbf{L}\mathbf{x}$ under the constraints $\mathbf{x} \in \{a, -b\}^N$, $a, b > 0$, $\|\mathbf{x}\| = 1$, and $\mathbf{x}^T\mathbf{1} = 0$ (the latter ensures that $\mathcal{V}_1$ and $\mathcal{V}_2$ are both nonempty). Spectral clustering is the relaxation of this problem in which the constraint $\mathbf{x} \in \{a, -b\}^N$ is dropped. The solution of the relaxed problem is given by the eigenvector $\mathbf{x} = \mathbf{u}_2$ of $\mathbf{L}$ corresponding to the second smallest eigenvalue λ_2; the cluster labels are then estimated as

$$\hat{l}_i = \begin{cases} 1, & x_i < 0, \\ 2, & x_i \geq 0. \end{cases} \tag{7}$$

2.2. Spectral Clustering on Signed Graphs

In signed graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, negative edge weights $W_{ij} < 0$ indicate dissimilarity of the adjacent nodes. This information is taken into account with the signed ratio cut (sRC)

$$\bar{\rho}(\mathcal{V}_1, \mathcal{V}_2) = \frac{2\gamma^+(\mathcal{V}_1, \mathcal{V}_2) + \gamma^-(\mathcal{V}_1) + \gamma^-(\mathcal{V}_2)}{|\mathcal{V}_1|\,|\mathcal{V}_2|}, \tag{8}$$

where $\gamma^+(\mathcal{V}_1, \mathcal{V}_2)$ is the positive weight of the cut set between $\mathcal{V}_1$ and $\mathcal{V}_2$,

$$\gamma^+(\mathcal{V}_1, \mathcal{V}_2) = \sum_{i \in \mathcal{V}_1} \sum_{j \in \mathcal{V}_2} \max\{0, W_{ij}\}, \tag{9}$$

and $\gamma^-(\mathcal{V}_k)$ is the aggregate negative weight within a cluster,

$$\gamma^-(\mathcal{V}_k) = -\sum_{i \in \mathcal{V}_k} \sum_{j \in \mathcal{V}_k} \min\{0, W_{ij}\}. \tag{10}$$

Accordingly, the sRC penalizes cuts of edges with large positive weights as well as intra-cluster edges with negative weights.

Similarly to the unsigned case, it can be shown that a relaxed version of the problem of minimizing the sRC amounts to minimizing the signed Laplacian form $\mathbf{x}^T\bar{\mathbf{L}}\mathbf{x}$ subject to the constraint $\|\mathbf{x}\| = 1$. Here, $\bar{\mathbf{L}}$ is the signed Laplacian matrix [19]

$$\bar{\mathbf{L}} = \bar{\mathbf{D}} - \mathbf{W}, \tag{11}$$

where

$$\bar{\mathbf{D}} = \operatorname{diag}\{\bar{d}_1, \ldots, \bar{d}_N\}, \qquad \bar{d}_i = \sum_{j=1}^{N} |W_{ij}|\,. \tag{12}$$

The signed Laplacian $\bar{\mathbf{L}}$ is positive semi-definite and even positive-definite for many real-world networks [19]. However, the constant vector $\mathbf{1}$ is no eigenvector of $\bar{\mathbf{L}}$. The solution of the signed SC problem is given the eigenvector $\mathbf{x} = \bar{\mathbf{u}}_1$ associated to the smallest eigenvalue $\bar{\lambda}_1$ of $\bar{\mathbf{L}}$. From this vector, the cluster labels are again obtained via (7).

3. SEMI-SUPERVISED SPECTRAL CLUSTERING

We next consider a semi-supervised scenario in which the true cluster labels l_i are known on a subset $\mathcal{S} = \{i_1, i_2, .., i_M\} \subseteq \mathcal{V}$ of nodes in the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$. This amounts to a graph signal processing problem aiming at reconstructing the binary label vector $\mathbf{l} \in \{1, 2\}^N$ from the known sample labels l_i, $i \in \mathcal{S}$. The idea is to incorporate the known label information in SC via a suitable modification of the edge weights.

In what follows, we assume that the edge weights are scaled such that $|W_{ij}| \leq 1$ (this can always be guaranteed by an appropriate weight rescaling that has no effect on SC). Let $\mathcal{S}_k \subset \mathcal{S}$ denote the set of nodes known to lie in cluster k, $k = 1, 2$ (i.e., $l_i = k$, $i \in \mathcal{S}_k$). We have $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. We

```
Algorithm 1 Semi-supervised signed SC
input: graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, sampled labels $l_i$, $i \in \mathcal{S}$
  construct the modified graph $\widetilde{\mathcal{G}}$ via (13)
  compute the signed Laplacian $\bar{\mathbf{L}}$ of $\widetilde{\mathcal{G}}$
  compute the first eigenvector $\mathbf{x}$ of $\bar{\mathbf{L}}$
  determine the cluster labels $\mathbf{l}$ via (7)
output: cluster labels $\mathbf{l}$
```

propose to construct a new signed graph $\widetilde{\mathcal{G}}$ from the existing graph $\mathcal{G}$ via the modified weight matrix $\widetilde{\mathbf{W}}$ with non-diagonal elements

$$\widetilde{W}_{ij} = \begin{cases} w_{\mathrm{sim}}, & (i,j) \in (\mathcal{S}_1 \times \mathcal{S}_1) \cup (\mathcal{S}_2 \times \mathcal{S}_2), \\ -w_{\mathrm{dis}}, & (i,j) \in (\mathcal{S}_1 \times \mathcal{S}_2) \cup (\mathcal{S}_2 \times \mathcal{S}_1), \\ W_{ij}, & \text{else.} \end{cases} \tag{13}$$

Here, $w_{\mathrm{sim}} > 0$ and $w_{\mathrm{dis}} > 0$ are suitably chosen weights (see below) for nodes in the same and in distinct clusters, respectively. The graph $\widetilde{\mathcal{G}}$ constructed in that manner has large positive edge weights $w_{\mathrm{sim}} > 0$ between nodes known to lie in the same cluster and large negative edge weights $-w_{\mathrm{dis}} < 0$ between nodes known to lie in distinct clusters. We note that for unsigned graphs, a similar idea with $w_{\mathrm{dis}} = 0$ was studied in [15].

Once the weight matrix $\widetilde{\mathbf{W}}$ of the modified graph has been constructed, we perform signed SC by computing the first eigenvector of the signed Laplacian $\bar{\mathbf{L}}(\widetilde{\mathcal{G}})$ and determining the cluster labels $\hat{l}_i$ according to (7). When performing SC, the weight matrix modifications promote cuts that correctly reflect the known cluster labels since

- new edges may be created;
- cuts across intra-cluster edges $(i,j) \in \mathcal{S}_k \times \mathcal{S}_k$ are penalized via w_{sim};
- cuts across inter-cluster edges $(i,j) \in \mathcal{S}_k \times \mathcal{S}_{3-k}$ are promoted via $-w_{\mathrm{dis}}$.

The overall proposed procedure for semi-supervised signed SC is stated in algorithm 1.

We next discuss the choice of the edge weights w_{sim} and w_{dis}. We would like to choose these weights such that the SC solution is consistent with the known sample labels, i.e., $\hat{l}_i = l_i$, $i \in \mathcal{S}$. Since the SC problem itself is not directly amenable to a corresponding analysis, we study the unrelaxed sRC minimization. By imposing

$$w_{\mathrm{dis}} > \frac{2w_{\mathrm{sim}}}{\max\{|\mathcal{S}_1|, |\mathcal{S}_2|\}}, \tag{14}$$

the minimum sRC for inconsistent label allocations (i.e., $\hat{l}_i \neq l_i$ for at least one $i \in \mathcal{S}$) is attained by the graph where all unmodified edge weights W_{ij} in (13) are zero with exactly one node from either $\mathcal{S}_1$ or $\mathcal{S}_2$ being placed in the wrong cluster and equals

$$\bar{\rho}(\mathcal{V}_1, \mathcal{V}_2) = \min_{k \in \{1,2\}} \left\{ \frac{2w_{\mathrm{sim}}(|\mathcal{S}_k| - 1) + w_{\mathrm{dis}}|\mathcal{S}_{3-k}|}{|\mathcal{V}_1|\,|\mathcal{V}_2|} \right\}. \tag{15}$$

This sRC is required to be larger than the maximum sRC of 2 for the case where all nodes in $\mathcal{S}$ are correctly labeled, attained by the fully connected graph with $W_{ij} = 1$ for all $i \neq j$. Since $|\mathcal{V}_1|\,|\mathcal{V}_2| \geq N^2/4$, a suitable choice of w_{sim} and w_{dis} is

$$\begin{aligned} w_{\mathrm{sim}} &= \frac{N^2}{8\,(\min\{|\mathcal{S}_1|, |\mathcal{S}_2|\} - 1)}, \\ w_{\mathrm{dis}} &= \frac{N^2}{4\max\{|\mathcal{S}_1|, |\mathcal{S}_2|\}}. \end{aligned} \tag{16}$$

For the unsigned case with $w_{\mathrm{dis}} = 0$, a similar line of arguments suggests to choose the intra-cluster edge weight such that $w_{\mathrm{sim}} \geq N^2/(4\min\{|\mathcal{S}_1|, |\mathcal{S}_2|\}$.

4. NUMERICAL EXPERIMENTS

We benchmark the performance of our proposed method (algorithm 1) against state-of-the art SC [15] and clustering by harmonic functions [20]. To this end, we performed Monte Carlo simulations with 10.000 realizations in two test scenarios with $N = 1000$ nodes. The performance metric is the number N_{err} of incorrectly labeled nodes. To ensure that the signed Laplacian has negative elements, the sampling set is chosen such that the first two samples are drawn one from each cluster and the remaining samples are drawn randomly among all nodes.

With the two-moon model (TM), two-dimensional data points are generated according to the model

$$\mathbf{y}_i = \begin{pmatrix} l_i - 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \cos(\varphi_i) \\ (3 - 2l_i)\sin(\varphi_i) \end{pmatrix} + \mathbf{n}_i. \tag{17}$$

Here, $l_i \in \{1, 2\}$ is the randomly drawn cluster label, $\varphi_i \sim \mathcal{U}(0, \pi)$ is a random angle, and $\mathbf{n}_i \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ is Gaussian noise. These data vector are noisy versions of points on the two unit-radius half-circles with centers $(0, 0)$ and $(1, 0)$ lying in the upper and lower halfplane, respectively. The graph representing this data is constructed by connecting each point to its five nearest neighbors (in Euclidean distance) and assigning the weights

$$W_{ij} = \exp\left(-\|\mathbf{x}_i - \mathbf{x}_j\|_2^2/2\right). \tag{18}$$

The second model is a random cluster graph (RCG), where two node clusters with $N/2$ nodes are generated with randomly placed edges such that two nodes in the same cluster are connected with probability $p_1 = 0.05$ and nodes in distinct clusters are connected with probability $p_2 = 0.02$ (all edges have weight 1).

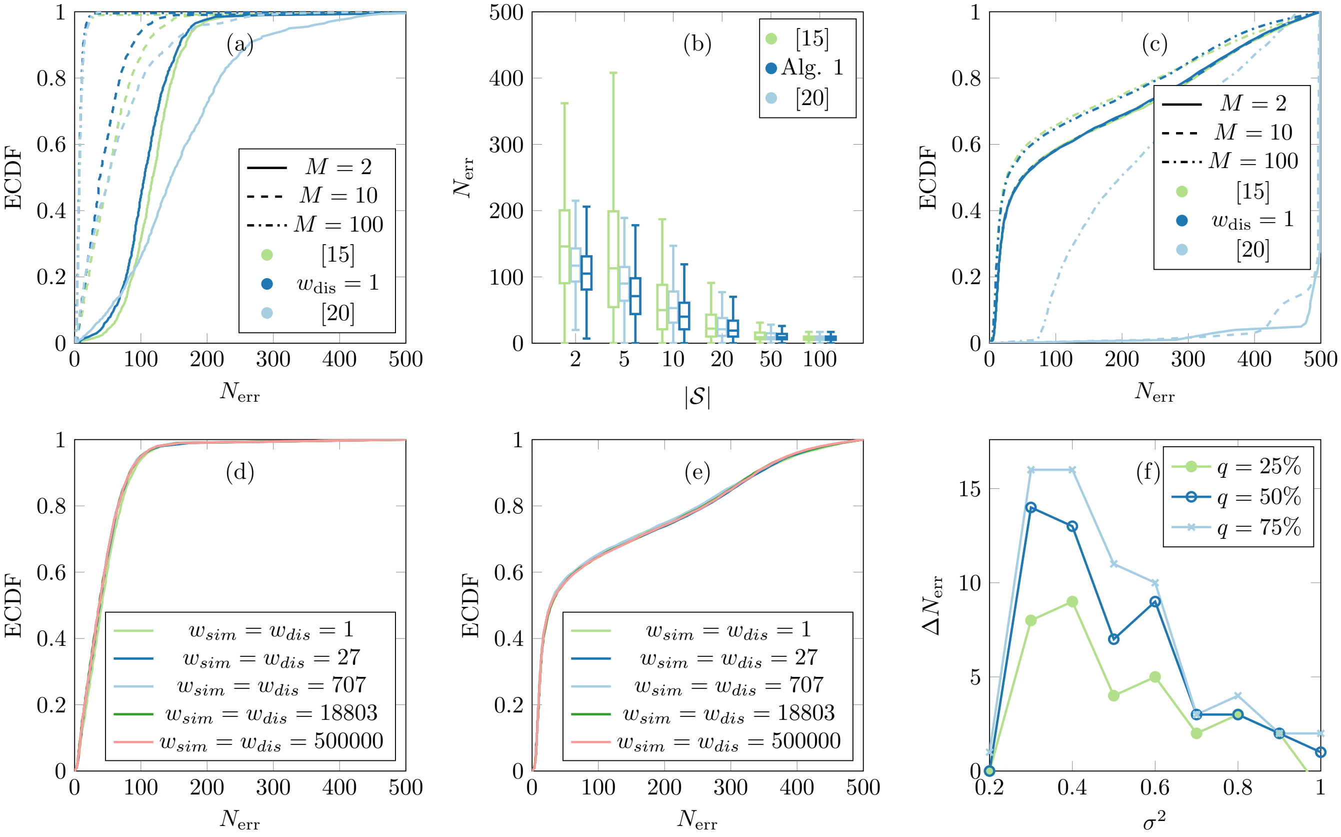

Fig. 1: Performance comparison of proposed schema and [15] and [20] (a) ECDF of N_{err} for the TM model with $\sigma^2 = 0.3$ and $w_{\mathrm{sim}} = w_{\mathrm{dis}} = 1$; (b) boxplot of N_{err} for the TM model with $\sigma^2 = 0.3$ for different M; (c) ECDF of N_{err} for the RCG model; (d) ECDF of N_{err} for the TM model with $\sigma^2 = 0.3$ with different edge weights $w_{\mathrm{sim}} = w_{\mathrm{dis}}$; (e) ECDF of N_{err} for the RCG model for different values of $w_{\mathrm{sim}} = w_{\mathrm{dis}}$; (f) clustering performance gap for $w_{\mathrm{sim}} = w_{\mathrm{dis}} = 1$ for different noise levels.

4.1. Performance for TM

We first compare the performance of our scheme using $w_{\mathrm{sim}} = w_{\mathrm{dis}} = 1$ with the methods from [15] and [20] for the TM model. Fig. 1(a) shows the ECDF of N_{err} for the three methods and $M = 2, 10, 100$. It is seen that our method performs better than the reference methods for both cases of 2 and 10 known labels. As expected, increasing the number of known labels improves performance for all schemes. Similar conclusions can be drawn from the box-plots for N_{err} and various M shown in Fig. 1(b).

For the RCG (see Fig. 1(c)) clustering is generally harder but the results are qualitatively similar, only that the advantage of our scheme here is less pronounced. The performance for $M = 2$ and $M = 10$ is almost identical since the number of samples here is too small to see a gain relative to unsupervised SC.

4.2. Impact of Edge Weights

We next analyze the performance of our scheme for different edge weights w_{sim} and w_{dis}. The ECDFs for the TM model and $M = 10$ samples and for the RCG model and $M = 100$ samples in Figs. 1(d) and (e), respectively. It can be seen that for both models distinct values of w_{sim} and w_{dis} have little influence on the clustering performance, with larger values performing slightly better than $w_{\mathrm{sim}} = w_{\mathrm{dis}} = 1$. This shows that the worst case edge weight choices (16) derived from a fully connected graph and a graph without edges (both of which have no real cluster structure) are overly pessimistic.

4.3. Impact of Noise Level

Finally, we investigate the impact of the noise variance σ^2 in the TM model that determines the amount of overlap between the two clusters. Figs. 1(f) shows the performance improvement of our scheme relative to [15], i.e. difference between the corresponding q-quantiles of the number N_{err} of incorrectly clustered nodes. It can be seen that for a very small noise variance, the performance gap is close to zero since here the clusters are almost perfectly separated. Similarly, for large noise variance the two moons strongly overlap and hence the cluster structure vanishes and both schemes fail.

For moderate σ^2, our proposed method performs uniformly better than [15].

5. CONCLUSION

In this paper we presented an extension of SC that modifies the weight matrix according to the knowledge about similarity or dissimilarity of the cluster labels of sampled nodes. For similar cluster labels the connecting edge is assigned a large positive weight and for dissimilar labels the edge weight is negative. We showed that SC using the signed Laplacian then results in a uniformly better performance than state-of-the art schemes in different graph models. We also found that the actual magnitude of the modified edge weights has little impact on performance.

REFERENCES

[1] M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” *PNAS*, vol. 99, no. 12, pp. 7821–7826, 2002.

[2] W. Liu, J. Wang, and S. F. Chang, “Robust and scalable graph-based semisupervised learning,” *Proc. IEEE*, vol. 100, no. 9, pp. 2624–2638, 2012.

[3] S. Fortunato, “Community detection in graphs,” *Physics Reports*, vol. 486, no. 3, pp. 75–174, 2010.

[4] S. E. Schaeffer, “Graph clustering,” *Computer Science Review*, vol. 1, no. 1, pp. 27 – 64, 2007.

[5] T. Jebara, J. Wang, and S. Chang, “Graph construction and b-matching for semi-supervised learning,” in *Proc. Int. Conf. Machine Learning*, Montreal, Quebec, Canada, June 2009, pp. 441–448.

[6] L. Hagen and A. B. Kahng, “New spectral methods for ratio cut partitioning and clustering,” *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 11, no. 9, pp. 1074–1085, 1992.

[7] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” in *Advances in Neural Information Processing Systems*, 2002, pp. 849–856.

[8] U. Von Luxburg, “A tutorial on spectral clustering,” *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[9] F. Wang and C. Zhang, “Label propagation through linear neighborhoods,” *IEEE Trans. on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 55–67, Jan. 2008.

[10] K. Ozaki, M. Shimbo, M. Komachi, and Y. Matsumoto, “Using the mutual k-nearest neighbor graphs for semi-supervised classification of natural language data,” in *Proc. Conf. Computational Natural Language Learning*, Portland, Oregon, USA, June 2011, pp. 154–162.

[11] P. Berger, M. Buchacher, G. Hannak, and G. Matz, “Graph learning based on total variation minimization,” in *in Proc. IEEE ICASSP*, Calgary, Alberta, Canada, Apr. 2018.

[12] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, “Learning Laplacian matrix in smooth graph signal representations,” *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, Dec. 2016.

[13] V. Kalofolias, “How to learn a graph from smooth signals,” in *Proc. Int. Conf. on Artificial Intelligence and Statistics*, vol. 51, Cadiz, Spain, May 2016, pp. 920–929.

[14] X. Y. Stella and J. Shi, “Multiclass spectral clustering,” in *Proc. IEEE Int. Conf. Computer Vision*, Oct 2003, pp. 313–319 vol.1.

[15] S. Kamvar, D. Klein, and C. Manning, “Spectral learning,” in *Int. Joint Conf. Artificial Intelligence*. Stanford InfoLab, 2003.

[16] X. Wang, B. Qian, and I. Davidson, “On constrained spectral clustering and its applications,” *Data Mining and Knowledge Discovery*, vol. 28, no. 1, pp. 1–30, 2014.

[17] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[18] A. Sandryhaila and J. M. F. Moura, “Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure,” *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sept. 2014.

[19] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, “Spectral analysis of signed graphs for clustering, prediction and visualization,” in *Proc. SIAM Int. Conf. Data Mining*. SIAM, 2010, pp. 559–570.

[20] X. Zhu, Z. Ghahramani, and J. D. Lafferty, “Semi-supervised learning using gaussian fields and harmonic functions,” in *Proc. Int. Conf. Machine learning (ICML-03)*, 2003, pp. 912–919.