

Identification of Cuts in Wireless Sensor Networks Scientific Through Spectral Classification

Suresh K R, Chandan Raj B

P.G. Student, Department of Computer Science & Engineering, EWIT, Bangalore , India¹

Associate Professor, Department of Computer Science & Engineering, Bangalore , India²

ABSTRACT: Wireless sensor networks (WSNs) often suffer from the disrupted connectivity due to unpredictable wireless channels, early depletion of node energy, and physical tampering by hostile users. The existence of a disconnected segment of the network referred to as network cut, leads to data loss, wasted power consumption, and congestion in the WSN. However, existing approaches to network cut detection in the WSN rely on the assumption that a node or a link either works normally or fails, without considering the uncertain and random features of wireless links in the WSN. In this paper, we extend the notion of the network cut based on the realistic wireless channel model. Furthermore, we formulate the problem of minimizing the normalized cut (Ncut) with critical nodes, considering the quality of wireless links, degree weights, and different priorities of sensor nodes. Then, we propose a network cut identification algorithm and dominant eigenvector computation algorithm that efficiently identify multiple network cuts by computing multiple Eigen values and eigenvectors according to a given parameter of eigen value gap. Extensive simulations are conducted to examine the effectiveness and robustness of the proposed approach. The results show that the proposed method strikes a balance between minimizing the Ncut objective and the degree of disconnection of critical nodes and achieves a better performance than existing algorithm.

I. INTRODUCTION

WIRELESS sensor networks (WSNs), consisting of large numbers of low power and low bandwidth wireless nodes, typically with unmanned and remote deployment, exhibit decentralized and self-organizing properties. The connectivity of a WSN can be easily disrupted due to fading wireless channels, early depletion of node energy, and physical tampering by hostile users. Furthermore, the channel fading and noise, the topology of a WSN and node failures will impact on network cut, i.e., a disconnected segment of the network. In particular, the network cut might cause data loss, wasted power consumption, and congestion problems. It is pointed out in [1] and [2] that the cut detection problem is potentially crucial in many WSN applications

Network Cut: Network cut has been studied extensively in graph theory and theoretical computer science. In the context of WSN, the network cut is typically caused by node failures, and such nodes are referred to as cut vertices. In [3]–[5], the nodes detect cut vertices locally but a spanning tree of all nodes across the network needs to be built first, which requires the overall network topology information, and incurs high communication overhead. In [6] and [7], distributed algorithms for the detection of cut vertices are developed to reduce the communication overhead.

In [8], a protocol is introduced to detect linear-shaped cuts in WSN by deploying a set of sentinels, which is a centralized algorithm, requiring the global topology information. A fully distributed algorithm is presented in [9] to detect arbitrarily shaped cuts. The cut detection algorithm in [10] is performed

on an overlay network consisting of a small number of representative nodes in order to speed up the convergence. Furthermore, in [11], a protocol is introduced that enables nodes to detect a cut with respect to multiple sink nodes and to detect a cut between any pair of source and destination. More generally, the maximum variation on the Weighted Spectrum [12] is proposed to evaluate geographically correlated vulnerable cuts or nodes in the network.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 10, May 2016

The above mentioned algorithms for detecting network cut are based on the assumption that each link or node has a binary state: normal or failure. On the other hand, the random nature of the wireless links calls for probabilistic modeling of the WSN.

II. SPECTRAL CLUSTERING

Our proposed approach to identifying network cuts in WSN is closely related to the minimum cut problem [19] of a graph, and spectral clustering [20] has become one of the most popular algorithms for solving the minimum cut problem, due to its simple implementation based on computing the eigenvectors and eigenvalues of the input matrix. Recent progresses on spectral clustering include the following. A graph-based semi-supervised learning algorithm [21] is used to provide the solution to the minimum cut problem based on given labeled data. A semi-supervised framework [22] is proposed for spectral clustering to improve the efficiency of the power method for computing the spectral clustering solution. Some common approaches to clustering problem based on pairwise distances are described in [23], and the theoretical guaranties for spectral clustering methods are provided. Two methods are proposed in [24] to efficiently combine the eigenvectors of the graph Laplacian matrices for studying the problem of clustering with data, and two algorithms are proposed in [25] to further improve the clustering quality for correlated probabilistic graphs. Our goal is to detect the fragility of WSN based on the realistic wireless channel model. Compared with the existing cut detection algorithms, our proposed algorithm can efficiently find multiple network cuts that have a high probability of occurrence, and separate a large number of nodes or critical nodes from the sink nodes? That is, our method can identify more network cuts that significantly affect the performance of WSN. Hence, after considering the quality of the underlying wireless links, the degree and different priorities of sensor nodes according to their locations and tasks, we propose a novel spectral clustering approach to identify network cuts in WSN. The contributions of this work are summarized as follows: 1) we extend the concept and properties of the network cut based on the realistic wireless channel model in the WSN, and define the problem of finding the network cut with the highest probability as a minimum cut problem. 2) We formulate the problem of minimizing normalized cut with critical nodes, which considers the degree and priority information of sensor nodes. A network cut identification (NCI) algorithm is developed to solve the above problem by computing the multiple eigenvectors of the regulated normalized graph Laplacian. 3) We propose a dominant eigenvector computation (DEC) algorithm to efficiently compute the multiple eigenvalues and eigenvectors according to a given gap parameter. Based on the Householder transformation, the Gershgorin theorem and the Sturm sequence property, the approximating eigenvalues can be obtained in order to accelerate the inverse power method for computing the corresponding eigenvectors. Notice that, although the solution that we provide here is centralized, no global topology information and no extra communication overhead are required before our algorithms are run. Each sensor node just communicates with its neighbor nodes to estimate the quality of wireless links. The estimation results are transmitted to sink nodes of the WSN. The sink nodes implement the NCI algorithm and the DEC algorithm to identify network cuts. Additionally, our approach can be applied to network management and network engineering in WSN. In fact, our algorithms can track the operational health of the infrastructure of a WSN, and ensure the connectivity between sensor nodes and sinks. This usually involves identifying network cuts that the network has suffered, and removing them by improving the quality of wireless communication between the nodes on the different sides of the network cuts, or by adding more sensor nodes. The remainder of this article is structured as follows. In Section 2, we discuss the system model and related background, and give the problem formulation. The details of our solution and algorithms are presented in Section 3, and the performance evaluation results are given in Section 4. We conclude in Section 5.

III. SYSTEM MODEL AND BACKGROUND

3.1 Network Cut Model

In this paper, a graph is used to represent a WSN. Let $G = (V, E)$ be an undirected graph with node set $V = \{v_1, v_n\}$ and link set $E = \{e_1, e_l\}$. We model the link failures as Bernoulli events: at any time, a link e_i is operational with probability p_i and it fails with probability $1 - p_i$. If the failures of links in a set $E_i = \{e_{i1}, e_{im}\}$ ($E_i \subset E$, $m < l$) can cause a subset

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 10, May 2016

of nodes to become disconnected from the sink nodes, then E_i is called a network cut. The probability of E_i failure is $P_{E_i} = \prod_{j=1}^m (1 - p_{ij})$. In general, a WSN is composed of a set of sensor nodes and multiple sink nodes, and sensor nodes assume different roles. We focus on the nodes that perform high priority tasks such as surveillance of the critical areas, or play an important role (e.g., high degree nodes). Such nodes are referred to as critical nodes.

We focus on the network cut E_i that will significantly affect the performance of WSN. In particular, E_i should satisfy the following properties: 1) E_i has a high p_{E_i} ; 2) E_i disconnects a large number of nodes from sink nodes; 3) E_i separates critical nodes from sink nodes.

TABLE 1
NETWORK CUT FOUND BY NCI ALGORITHM

From the above discussion, network cuts are closely related to graph partitioning. It is known that spectral clustering can be used to solve the relaxed versions of graph partitioning

problems [20].

3.2 NETWORK CUT IDENTIFICATION ALGORITHM.

Spectral Clustering Based on Rayleigh-Ritz Theorem Here we aim to solve problem (11) based on the Rayleigh-Ritz theorem. To that end, we first verify the following three properties: 1) The objective function in (11) can be rewritten in the same form as (8), that is $f^T H_0 f$. 2) H_0 , defined as the normalized graph Laplacian of $G = (V, E)$, can be expressed as $D^{-1/2} W_0 D^{-1/2}$, and $j(W_0)_{ij} = d_i$, where $d_i = \sum_k D_{ii} = \sum_k W_{ik}$. Then we get $f^T L_0 f = \frac{1}{2} N_{cut}(V_1, V_2) \text{vol}(V)$, where $L_0 = D - W_0$. Furthermore, we have $1 - (f^T H_0 f) / (f^T f) = f^T L_0 f / \text{vol}(V) = \frac{1}{2} N_{cut}(V_1, V_2)$, where $f = D^{-1/2} h$. So minimizing N_{cut} is equivalent to maximizing the Rayleigh quotient. 3) The eigenvectors of H_0 as solutions to (11) should be orthogonal to $D^{-1/2} e$.

Algorithm 1 Network Cut Identification

Input: Given an undirected graph $G = (V, E)$ with n nodes

- 1: Sink nodes collect the link quality information in the WSN, calculate the adjacency matrix W whose elements correspond to $\omega_{ij} = (-\log(1 - p_{ij})) \geq 0$ of edge e_{ij} .
- 2: Define the critical node set A_2 and the sink node set A_1 to obtain $\phi = (\phi_1, \phi_n)^T \in R^n$ given by (9)-(10).
- 3: Calculate the diagonal degree matrix D , $H = D^{-1/2} W D^{-1/2}$, and $H_0 = H + (1-\gamma)\phi\phi^T$, where γ is an adjustable parameter.
- 4: Given an eigenvalue gap parameter $0 < \eta < 1$, call the dominant eigenvector computation (DEC) algorithm (Algorithm 2) to obtain the eigenpair set $\{(f_i(H_0), \lambda_i(H_0)), 1 \leq i \leq m < n\}$ such that $\lambda_{i+1}(H_0) - \lambda_i(H_0) \geq \eta$ for all $1 \leq i \leq m - 1$ and $\lambda_{m+1}(H_0) - \lambda_m(H_0) < \eta$.
- 5: By taking the sign of $f_i(H_0)$ ($1 \leq i \leq m$), we obtain m network cut solutions to (11).

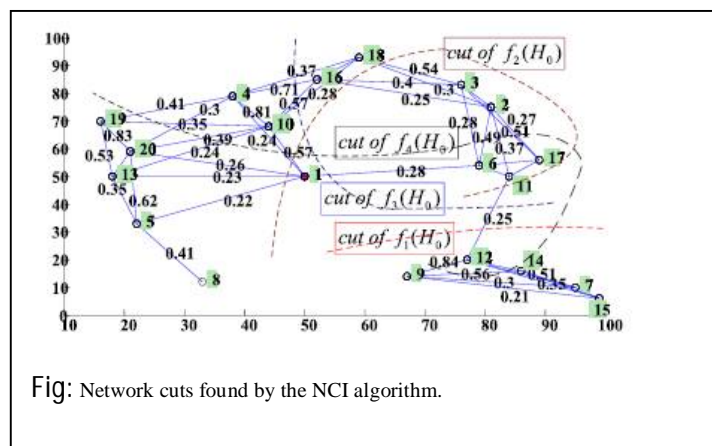


Fig: Network cuts found by the NCI algorithm.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 10, May 2016

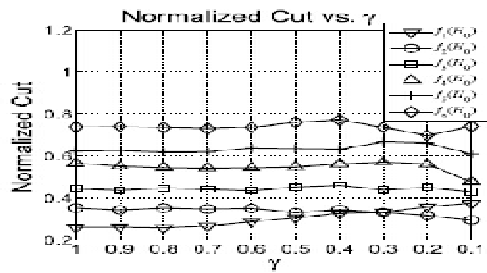


Fig Impact of γ on the normalized cut N cut

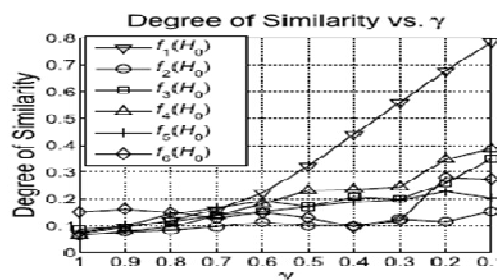
In order to make our description more clear, we use a network example to show how the NCI algorithm works. As shown in Fig. There are 20 nodes randomly spread over 100×100 sensor field, and the weighted links are shown in the figure. The node labelled by '1' is chosen as the sink Node we refer to the nodes that are directly connected to the Sink node as critical nodes. We perform the NCI Algorithm using $\gamma = 0.5$ in (11) and the eigenvalue gap $\eta = 0.8$ to find the network cuts. The corresponding results are summarized in Table I.

IV. SIMULATION RESULTS

As discussed in Section II.A, each link $ek \in E$ carries a weight $\omega_k = -\log(1 - pk) \geq 0$. The weight ω_k , as a function of the link probability pk , is not directly observed and can only be estimated through link quality measurements. At the same time, the packet reception ratio (PRR), being an important QoS metric, is used to estimate the quality of the underlying Wireless links in the WSN [16]. In practice, the PRR can be obtained by computing the ratio of the number of received Packets to the number of sent packets. Therefore, we use the PRR of link $ek \in E$ as the link weight instead of ω_k .

Eigenvalue	V_1	V_1
$\lambda_1(H_0) = 0.9797$	1 2 3 4 5 6 8 10 11 13 16 17 18 19 20	7 9 12 14 15
$\lambda_2(H_0) = 0.8800$	1 2 3 6 11 17	4 5 7 8 9 10 12 13 14 15 16 18 19 20
$\lambda_3(H_0) = 0.8044$	1 4 5 7 8 9 10 12 13 14 15 19 20	2 3 6 11 16 17 18
$\lambda_4(H_0) = 0.6533$	1 5 6 8 11 12 13 17 19 20	2 3 4 7 9 10 14 15 16 18

In our setup, 100 sensor nodes are uniformly distributed in a $200m \times 200m$ square region. We use the log-normal shadowing model to characterize the wireless signal propagation



International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 10, May 2016

Fig. shows the N_{cut} values corresponding to the multiple

Solutions of $f_1(H_0)$. . . $f_6(H_0)$ found by the NCI algorithms, corresponding to eigenvalues $1 \geq \lambda_1(H_0) \geq \dots \geq \lambda_6(H_0)$. We can observe that $f_1(H_0)$ can provide the best clustering performance, and the performance degrades going from $f_1(H_0)$ to $f_6(H_0)$.

Recall that the parameter γ determines the tradeoff between the degree of similarity and the N_{cut} objective. Larger γ leads to the better clustering performance, whereas smaller γ leads to a higher degree of similarity between $f_i(H_0)$ and ϕ . Hence, As γ is reduced, NCI is inclined to obtain the cut solution that disconnects more critical nodes from the sink nodes. In Fig. , We can observe the N_{cut} value corresponding to $f_1(H_0)$

Increases due to the clustering performance degradation, Whereas the degree of similarity corresponding to $f_1(H_0)$ Increases rapidly in Fig. . Furthermore, both the degree of Similarity and the N_{cut} corresponding to the other $f_i(H_0)$ ($i = 2, \dots, 6$) are nonsensitive to the choice of γ . Based on the above analysis, in the sequel, only $f_1(H_0)$ is Chosen to evaluate the performance of the NCI algorithm.

Impact of Wireless Channel Parameters

In this set of simulations, we study the impact of the wireless channel parameters on the performance of the NCI algorithm. First we consider the effect of the path loss exponent β that measures the rate at which the received signal strength decreases with distance. From the analysis of (24), we know the received signal Strength becomes weaker with the increase of the path loss exponent β when the distance remains constant. Hence, the Value of $p_{rr}(d)$ decreases accordingly. That is, the NCI Algorithm can obtain better clustering performance as β increases. That the N_{cut} decreases as β increases from 1 to 2.5. Furthermore, the clustering performance corresponding to $\gamma = 0.1, \gamma = 0.5$ and $\gamma = 0.9$ is in ascending order. In Fig. we can observe that the NCI algorithm with the smaller γ is more likely to obtain better critical nodes matching performance, and the degree of similarity for $\gamma = 0.1$ is insensitive to the value of β due to the fact that β can significantly affect the normalized cut rather than the degree of similarity. Moreover, when $\beta < 2$, the degree of similarity for $\gamma = 0.5$ decreases with β due to the fact that the NCI algorithm can trade off between the clustering performance and the critical nodes matching performance. Hence, the smaller N_{cut} leads to the lower degree of similarity. However, as β increases, more and more solutions can be chosen according to a given eigenvalue gap η due to the connection quality of WSN getting worse, consideration can be given to both the clustering and the matching performance, so the degree of similarity for $\gamma = 0.5$ increases when $\beta > 2$. The degree of similarity for $\gamma = 0.9$ behaves similarly as

for $\gamma = 0.5$. Recall that we can use the standard deviation σ_{dB} of X_{dB} to capture the time-varying nature of the signal due to the different surroundings on the propagation path. From (26), since the Q-function is a nonotonically decreasing function, $p_{rr}(d)$ increases with σ_{dB} , and the NCI algorithm obtains worse clustering performance accordingly. Fig. shows when σ_{dB} varies from 2 to 9, the N_{cut} increases gradually, and the clustering performance corresponding to $\gamma = 0.1, \gamma = 0.5$ and $\gamma = 0.9$ is in ascending order.

Robustness of the NCI Algorithm

In practice, there exist estimation errors in the link probabilities, we next study how the performance of the proposed algorithm is affected if we add a perturbation to the adjacency matrix, that is $\hat{W} = W +$, where contains additive White Gaussian noise (AWGN) with the specified perturbation level (PL), where the PL is the $p_{rr}(d)$ to estimation noise ratio in dB, and we set $PL = 30dB$. In the clustering performance corresponding to $\gamma = 0.1, \gamma = 0.5$ and $\gamma = 0.9$ is in ascending order, and the N_{cuts} for three different γ decrease as the PL increases from 5dB to 40dB due to decreasing estimation error. There exists an SNR threshold such that the N_{cut} remains almost constant when PL is larger than the threshold, and the N_{cut} decrease rapidly with the PL when the PL is less than the threshold. can also observe that the threshold corresponding to $\gamma = 0.9$ (about 10 dB) is smaller than the thresholds corresponding to the other two cases. Hence, we can conclude that the NCI algorithm corresponding to larger γ achieves the better robustness performance. Moreover, In , the Critical nodes matching performance corresponding to $\gamma = 0.1, \gamma = 0.5$ and $\gamma = 0.9$ is in descending order. When the N_{cut} decreases with PL, the degree of similarity decreases accordingly due to the trade-off between the clustering performance and the matching performance.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 10, May 2016

V. CONCLUSION

In this article, we aim to identify network cuts in the WSN, While taking fully into account the quality of wireless links, degree weights and different priorities of sensor nodes. Based on the spectral clustering theory, we have proposed a scheme based on minimizing the normalized cut with critical nodes. We have further developed the network cut identification (NCI) algorithm and the dominant eigenvector computation (DEC) algorithm, which provide the efficient solution by computing the multiple eigenvalues and eigenvectors according to a given parameter of eigenvalue gap, requiring no global topology information. Simulation results have demonstrated the effectiveness and robustness of the proposed algorithm decreases with PL, the degree of similarity decreases accordingly due to the trade-off between the clustering performance and the matching performance.

REFERENCES

- [1] C.-Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.
- [2] A. Cerpa and D. Estrin, "ASCENT: Adaptive self-configuring sensor networks topologies," *IEEE Trans. Mobile Comput.*, vol. 3, no. 3, pp. 272–285, Jul./Aug. 2004.
- [3] J. Cheriyan, M.-Y. Kao, and R. Thurimella, "Scan-first search and sparse certificates: An improved parallel algorithm for k -vertex connectivity," *SIAM J. Comput.*, vol. 22, no. 1, pp. 157–174, 1993.
- [4] W. Hohberg, "How to find biconnected components in distributed networks," *J. Parallel Distrib. Comput.* vol. 9, no. 4, pp. 374–386, 1990.
- [5] S. T. Huang, "A new distributed algorithm for the biconnectivity problem," in *Proc. Int. Conf. Parallel Process.*, Dupage County, IL, USA, 1989, pp. 106–113.
- [6] I. Stojmenovic, D. Simplot-Ryl, and A. Nayak, "Toward scalable cut vertex and link detection with applications in wireless ad hoc networks," *IEEE Netw.*, vol. 25, no. 1, pp. 44–48, Jan./Feb. 2011.
- [7] X. Liu, X. Li, and A. Kreling, "A fully distributed method to detect and reduce cut vertices in large-scale overlay networks," *IEEE Trans. Comput.*, vol. 61, no. 7, pp. 969–985, Jul. 2012.
- [8] N. Shrivastava, S. Suri, and C. D. Toth, "Detecting cuts in sensor networks," *ACM Trans. Sensor Netw.*, vol. 4, no. 2, pp. 1–25, 2008.
- [9] P. Barooah, H. Chenji, R. Stoleru, and T. Kalmar-Nagy, "Cut detection in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 3, pp. 483–490, Mar. 2012.
- [10] M. Won, S. M. George, and R. Stoleru, "Towards robustness and energy efficiency of cut detection in wireless sensor networks," *Ad Hoc Netw.*, vol. 9, no. 3, pp. 249–264, 2011.
- [11] M. Won and R. Stoleru, "Destination-based cut detection in wireless sensor networks," in *Proc. IFIP 9th Int. Conf. Embedded Ubiquitous Comput. (EUC)*, Melbourne, Australia, Oct. 2011, pp. 55–62.
- [12] X. Long, D. Tipper, and T. Gomes, "Evaluating geographic topological vulnerabilities," in *Proc. 5th Int. Congr. Ultra Modern Telecommun. Control Syst. Workshops (ICUMT)*, Almaty, Kazakhstan, 2013, pp. 54–61.
- [13] S. Kar and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3315–3326, Jul. 2008.
- [14] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.
- [15] R. R. Brooks, B. Pillai, S. Racunas, and S. Rai, "Mobile network analysis using probabilistic connectivity matrices," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 4, pp. 694–702, Jul. 2007.
- [16] N. Pustchi and T. Korkmaz, "Improving packet reception rate for mobile sinks in wireless sensor networks," in *Proc. IEEE Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, San Francisco, CA, USA, Jun. 2012, pp. 1–9.
- [17] X. Ma, "Packet reception ratios in two-dimensional broadcast ad hoc networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Maui, HI, USA, Jan. 2012, pp. 412–416.
- [18] L. Wang, C. Goursaud, N. Nikaen, L. Cottatellucci, and J. Gorce, "Cooperative scheduling for coexisting body area networks," *IEEE Trans. Wireless Commun.* vol. 12, no. 1, pp. 123–133, Jan. 2013.
- [19] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Math. J.*, vol. 23, no. 98, pp. 298–305, 1973.
- [20] U. V. Luxburg, "A tutorial on spectral clustering," *Statist. Comput.* vol. 17, no. 4, pp. 395–416, 2007.
- [21] X. Zhu and A. B. Goldberg, *Introduction to Semi-Supervised Learning*, San Mateo, CA, USA: Morgan Kaufmann, 2009.
- [22] D. Mavroudis, "Accelerating spectral clustering with partial supervision," *Data Mining Knowl. Discovery*, vol. 21, no. 2, pp. 241–258, 2010.
- [23] E. Arias-Castro, "Clustering based on pairwise distances when the data is of mixed dimensions," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1692–1706, Mar. 2011.
- [24] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov, "Clustering with multi-layer graphs: A spectral perspective," *IEEE Trans. Signal Process.* vol. 60, no. 11, pp. 5820–5831, Nov. 2012.
- [25] Y. GU, C. GAO, G. Cong, and G. Yu, "Effective and efficient clustering methods for correlated probabilistic graphs," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 5, pp. 1117–1130, May 2014.
- [26] H. Lutkepohl, *Handbook of Matrices*. Chichester, U.K.: Wiley, 1997.
- [27] C. R. Hadlock, *Field Theory and Its Classical Problems, the Carus Mathematical Monographs*. Washington, DC, USA: MAA, 1978.
- [28] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.
- [29] T. S. Rappaport, *Wireless communications: Principles and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999