CISTER
Research Centre in
Real-Time & Embedded
Computing Systems

Journal Paper

Multi-Gateway Designation for Real-Time TSCH Networks using Spectral Clustering and Centrality

Miguel Gutiérrez Gaitán*
Diego Dujovne
Julián Zuñiga
Alejandro Figueroa
Luís Almeida*

*CISTER Research Centre

CISTER-TR-220904

2022

Multi-Gateway Designation for Real-Time TSCH Networks using Spectral Clustering and Centrality

Miguel Gutiérrez Gaitán*, Diego Dujovne, Julián Zuñiga, Alejandro Figueroa, Luís Almeida*

*CISTER Research Centre

Rua Dr. António Bernardino de Almeida, 431

4200-072 Porto

Portugal

Tel.: +351.22.8340509, Fax: +351.22.8321159

E-mail: mjggt@isep.ipp.pt, diego.dujovne@mail.udp.cl, julian.zuniga1@mail.udp.cl, alejandro.figueroa@unab.cl, lda@fe.up.pt

https://www.cister-labs.pt

Abstract

This letter proposes a multi-gateway designation framework to design real-time wireless sensor networks (WSNs) improving traffic schedulability, i.e., meeting the traffic time constraints. To this end, we resort to Spectral Clustering un-supervised learning that allows defining arbitrary k disjoint clusters without knowledge of the nodes physical position. In each cluster we use a centrality metric from social sciences to designate one gateway. This novel combination is applied to a time-synchronized channel-hopping (TSCH) WSN under earliest-deadline-first (EDF) scheduling and shortest-path routing. Simulation results under varying configurations show that our framework is able to produce WSN designs that greatly reduce the worst-case network demand. In a situation with 5gateways, 99% schedulability can be achieved with 3.5 times more real-time flows than in a random benchmark.

Multi-Gateway Designation for Real-Time TSCH Networks using Spectral Clustering and Centrality

Miguel Gutiérrez Gaitán, Diego Dujovne, Julián Zuñiga, Alejandro Figueroa and Luís Almeida

***Abstract*—This letter proposes a multi-gateway designation framework to design real-time wireless sensor networks (WSNs) improving traffic schedulability, i.e., meeting the traffic time constraints. To this end, we resort to Spectral Clustering unsupervised learning that allows defining arbitrary k disjoint clusters without knowledge of the nodes physical position. In each cluster we use a centrality metric from social sciences to designate one gateway. This novel combination is applied to a time-synchronized channel-hopping (TSCH) WSN under earliest-deadline-first (EDF) scheduling and shortest-path routing. Simulation results under varying configurations show that our framework is able to produce WSN designs that greatly reduce the worst-case network demand. In a situation with 5 gateways, 99% schedulability can be achieved with 3.5 times more real-time flows than in a random benchmark.**

***Index Terms*—Centrality, Clustering, EDF, TSCH, WSN.**

I. Introduction

THE Industrial Internet of Things (IIoT) revolutionized factories with a remarkable increase of wireless technologies [1]. For example, WirelessHART, ISA100.11a and 6TiSCH [2] became popular in industrial monitoring and process control [3]. TSCH or time-slotted channel-hopping is a medium access control (MAC) layer common to these technologies offering real-time communication features.

In the context of IIoT, many networks are WSNs with a fixed set of field nodes gathering data from a process to support system wide monitoring and control. The network flows are static, defined at design time, and converge to one or more gateways. Typically, gateways are deployed in some arbitrary positions. Conversely, gateway designation uses existing nodes for the gateway function without adding new ones, but their positions are constrained, instead.

Proper gateway designation, routing and scheduling are needed for real-time flows to meet their time constraints, e.g., deadlines. Recent work showed that *network centrality* (a concept from social networks) can support gateway designation to significantly improve the real-time performance of TSCH WSNs by design [4]. Prior studies also showed that using multiple gateways (a.k.a. sinks) increases parallelization of flows and improves the traffic timeliness in industrial WSNs [5]. However, to our best knowledge, none of these works used *schedulability*, i.e. ability to satisfy all traffic time constraints, as a primary design criterion.

Thus, our research contributes to the state-of-the-art in real-time TSCH WSNs providing a multi-gateway designation framework for improved schedulability that consists on first clustering the network with Spectral Clustering [6], which does not require knowledge of nodes positions, just adjacency, and then applying a centrality metric to designate a gateway inside each cluster. Note that clustering is used only for gateway designation, being orthogonal to traffic scheduling and routing that operate globally across the network. Moreover, our gateway designation framework is parametric, allowing exploring the trade-off between schedulability and k, the number of gateways, thus supporting efficient designs with the minimum number of gateways needed to guarantee traffic schedulability. Finally, simulation results with random topologies and message sets achieve 99% schedulability using 5 gateways with 3.5 times more real-time flows than a random gateway designation.

Manuscript received June 8, 2022. Revised August 7, 2022. Accepted September 13, 2022. Date of publication October XX, 2022. This work was co-financed by national funds through FCT/MCTES (Portuguese Foundation for Science and Technology) within the CISTER Research Unit (UIDB/04234/2020); by the Operational Programme for Competitiveness and Internationalization (COMPETE 2020) under the PT2020 Agreement, through the European Regional Development Fund (ERDF); and also by FCT and the ESF (European Social Fund) through the Regional Operational Programme (ROP) Norte 2020, under PhD grant 2020.06685.BD. The authors also thank to the CYTED AgIoT Project (520RT011), CORFO COTH2O "Consorcio de Gestión de Recursos Hídricos para la Macrozona Centro-Sur" and Proyecto Asociativo UDP "Plataformas Digitales como Modelo Organizacional".

Miguel G. Gaitán and Alejandro Figueroa are with the Andrés Bello University, Santiago, Chile. E-mail: {miguel.gutierrez, alejandro.figueroa}@unab.cl.

Miguel G. Gaitán and Luís Almeida are with the Research Centre in Real-Time and Embedded Computing Systems (CISTER), and with the University of Porto, Porto, Portugal. E-mail: {mgg, lda}@fe.up.pt.

Diego Dujovne and Julián Zuñiga are with the Diego Portales University, Santiago, Chile. Email: {diego.dujovne,julian.zuniga1}@mail.udp.cl.

II. Related Work

Chen et al. [7] studied the minimum sink placement in TSCH networks with latency and reliability guarantees, and showed the problem is NP-hard. They proposed an algorithm to solve it by jointly considering the RPL routing protocol and DeTAS scheduling, but not considering schedulability analysis. Dobslaw et al. [5] addressed *schedulability* as a QoS constraint by proposing a complete cross-layer configuration for industrial WSN that considers, among other, the possibility of adding multiple sinks. Their work also used clustering, but based on k-*means* that requires actual nodes positions, thus being not applicable to WSNs represented by connectivity graphs and lacking the position of nodes as in our case. Conversely, *spectral clustering* [6], which we use, can be directly applied to connectivity graphs, e.g., expressed as adjacency matrices. Several other studies in the literature have addressed alike problems (see e.g. [8]–[10]), either from the perspective of clustering and/or from the viewpoint of multi-sink placement, targeting common delay or reliability issues.

However, to the best of our knowledge, none provided schedulability analysis and commonly require knowledge of the nodes positions and use an arbitrary gateway placement approach. Our work is the first to provide a position-agnostic multi-gateway designation method for improved network *schedulability*. For this reason, we compare against a random gateway designation baseline, similarly to other works addressing the single gateway designation problem [11].

III. System Model

A. Network Model

A WSN is abstracted using an undirected graph $G = (V, E)$, where V and E are the set of vertices (nodes) and edges (links), respectively. Nodes can perform the sensing, relaying or gateway functions and are connected wirelessly forming a mesh. The total number of nodes is $N = |V|$ among which k are gateways.

The network is assumed to be TSCH-based, thus globally synchronized using a TDMA framework with multiple channels (up to $m = 16$). This enables multi-hop concurrent and single-packet transmissions at each slot/hop. All transmissions are assumed to be triggered following a global traffic schedule. We assume this plan is made according to an EDF scheduling policy and shortest-path source routing.

B. Flow Model

We consider a number of n sensor nodes transmitting periodically their sensing data toward any of the k gateways. These messages are required to reach the gateways before specific timing constraints (deadlines). We denote as $F = \{f_1, f_2, \ldots, f_n\}$ the set n of time-sensitive flows. Each flow is characterized by a tuple (C_i, D_i, T_i, ϕ_i), where C_i is the transmission time between the node s_i and one of the k gateways, T_i is the transmission period, D_i is the (relative) deadline, and ϕ_i is the multi-hop routing path. Note that each flow releases potentially an infinite number of transmissions. Formally, the γ^{th} instance is released at time $r_{i,\gamma}$ such that $r_{i,\gamma+1} - r_{i,\gamma} = T_i$. Then, by following the EDF policy, $d_{i,\gamma} = r_{i,\gamma} + D_i$ denotes the absolute deadline for $f_{i,\gamma}$ to arrive at its gateway (destination).

C. Performance Model

We use the supply/demand-bound based schedulability analysis in [12] to evaluate the ability of the WSN to satisfy the timing constraints of all flows in the network as in [13]. Specifically, this method checks whether the *supply-bound function* (sbf) is greater than or equal to the so-called *forced-forward demand-bound function* (FF-DBF) adapted for WSNs. The sbf is the minimal transmission capacity offered by a WSN with m channels, while FF-DBF is the upper bound on the demand generated by a set F of n deadline-constrained flows when evaluated in any interval of length ℓ.

Eq. 1 formally presents the supply/demand-bound based schedulability test where $\mathrm{sbf}(\ell)$ fulfils the conditions in Eq. 2 and the FF-DBF-WSN is defined by Eq. 3.

$$\text{FF-DBF-WSN}(\ell) \leq \mathrm{sbf}(\ell),\ \forall \ell \geq 0 \tag{1}$$

$$\mathrm{sbf}(0) = 0 \land \mathrm{sbf}(\ell + h) - \mathrm{sbf}(\ell) \leq m \times h, \forall \ell, h \geq 0 \tag{2}$$

$$\text{FF-DBF-WSN}(\ell) = \overbrace{\frac{1}{m}\sum_{i=1}^{n} \text{FF-DBF}(f_i, \ell)}^{\text{CHANNEL CONTENTION}} + \underbrace{\sum_{i,j=1}^{n}\left(\Delta_{i,j} \cdot \max\left\{\left\lceil \frac{\ell}{T_i}\right\rceil, \left\lceil \frac{\ell}{T_j}\right\rceil\right\}\right)}_{\text{TRANSMISSION CONFLICTS}} \tag{3}$$

Note that FF-DBF-WSN results from two terms contributing to the worst-case network demand: $i)$ *channel contention* and $ii)$ *transmission conflicts*. The former, expressed by the left parcel of Eq. 3, is equivalent to FF-DBF for multiprocessors [14], here used to model the mutually exclusive condition for scheduling on multiple channels. The latter, expressed by the right parcel, represents the delay contribution due to multiple flows converging on a common half-duplex node. Eq. 4 presents $\Delta_{i,j}$ as the overlapping factor between the the paths of flows f_i and $f_j \in F$ (with $i \neq j$) as defined in [15].

$$\Delta_{i,j} = \sum_{q=1}^{\delta(ij)} Len_q(ij) - \sum_{h'=1}^{\delta'(ij)} (Len_{q'}(ij) - 3) \tag{4}$$

$\delta(ij)$ is the total number of overlaps between f_i and f_j of which $\delta'(ij)$ are the ones larger than 3. $Len_q(ij)$ and $Len_{q'}(ij)$ are the respective q and q' overlap lengths between f_i and f_j, with $h \in [1, \delta(ij)]$ and $q' \in [1, \delta'(ij)]$. After 3 hops, slots can be reused, not causing further transmission conflicts.

IV. Clustering-Assisted Multi-Gateway Designation for real-time WSNs

Given the network, flow and performance models presented in Section III, we consider the problem of how to judiciously designate multiple nodes as gateways (or sinks) to enhance network schedulability in real-time WSNs. We propose a framework combining *spectral clustering*, which is conveniently position-agnostic, with *centrality* metrics. By doing this, we also generalize to arbitrary gateways the single gateway method in [4]. We recall that our framework is intended to be used at system design time, assuming full knowledge of the network topology (graph) in the form of an adjacency matrix representing binary connectivity with lossless links.

A. Spectral Clustering

We built upon *spectral clustering* [6] to virtually partition the network G in a set of k disjoint and arbitrarily shaped clusters. The key idea of the method is to leverage the eigen-decomposition of the graph Laplacian matrix (L) to find solutions based on the relaxation of graph cut problems. In this work, we use the direct k-way spectral clustering algorithm proposed by Ng, Jordan and Weiss [16] to identify groups of widely separated nodes represented by k connected subgraphs.

Differently from other spectral clustering methods, the Ng-Jordan-Weiss (NJW) algorithm uses eigenvectors of the normalized Laplacian (L_{norm}) which can be computed as follows:

$$L_{norm} = D^{-1/2} \cdot L \cdot D^{-1/2} \tag{5}$$

where D is the degree matrix, i.e., the diagonal matrix with the degrees of the nodes, and $L = D - A$ is the Laplacian, with

Algorithm 1 NJW Spectral Clustering [16]

Input: a graph G and the target number of clusters k
Output: a partition of k clusters $\Pi = \{G_1, G_2, ..., G_k\}$

1: Find the first k eigenvectors u_1,u_2,...,u_k of L_{norm} and sort them in the columns of U'
2: Build matrix $U = [u_{ij}]_{n \times k}$ based on U' by normalizing each row of U' using $u_{ij} = u'_{ij}/\sqrt{\sum_k u'^2_{ik}}$
3: Let the i^{th} row of the matrix U represent node v_i from graph G
4: Apply k-means algorithm (or an equivalent method) to U and find a k-way partitioning $\Pi' = \{G'_1, ..., G'_n\}$
5: Form the final partition Π assigning every node v_i to the cluster G_ℓ, if the i^{th} row of U belongs to G'_ℓ in Π'

A being the adjacency matrix of the graph. For completeness, we revisit the NJW spectral clustering method in Algorithm 1. Though step 4 uses *k-means*, it applies virtual distances, after the normalization in step 2, thus keeping the independence from physical positions.

B. Centrality-Driven Multi-Gateway Designation

After a number of k clusters has been identified, a *network centrality* metric is applied per cluster to designate the respective gateway as the node in the cluster with the highest centrality measure. We consider the four most common metrics in social network analysis, namely, $i)$ degree, $ii)$ betweenness, $iii)$ closeness, and $iv)$ eigenvector centrality. These are considered near optimally correlated for the purposes of benchmarking [17]. These definitions are now taken with the interpretation of being *cluster centrality* metrics, where each cluster G_ℓ is a subgraph of G and is characterized by a cluster adjacency matrix A_ℓ and a number of nodes N_ℓ, with $\sum_{l=1}^{k} N_\ell = N$. Table I summarizes the formal expressions of the four *cluster centrality*[1] metrics when applied to a given node v_q of cluster G_ℓ, where $q \in [1, N_\ell]$.

TABLE I: Cluster Centrality Metrics.

Metric	Definition
Degree	$DC(v_q) = \frac{degree(v_q)}{N_\ell - 1}$
Betweenness	$BC(v_q) = \sum_{q \neq r} \frac{sp_{r,s}(v_q)}{sp_{r,s}}$
Closeness	$CC(v_q) = \frac{1}{\sum_{p \neq q} distance(v_p, v_q)}$
Eigenvector	$EC(v_q) = \frac{1}{\lambda_{max}(A_k)} \cdot \sum_{j=1}^{N_k} a_{j,q} \cdot x_j$

V. Performance Evaluation

A. Simulation Setup

*1) **Network topologies**:* We consider 1000 random topologies built upon the synthetic generation of network graphs. Each topology is generated with a target node density $d = 0.1$ using a sparse uniformly distributed binary random matrix of $N \times N$, i.e. assuming lossless links, where N is the total number of network nodes, including the k gateways. Without loss of generality, we use $k = \{1, 3, 5\}$ and $N = 75$ for all the simulation experiments, similarly to the validation carried out in [9].

*2) **Gateway designation**:* After clusters have been created, a number of k nodes is selected as gateways using each of the cluster centrality metrics in Table I. In the absence of a better solution to compare against, a random designation of k gateways is also considered, for benchmarking. The random designation is done before clusters are created.

*3) **Network flows**:* A subset of $n \in [1, 30]$ vertices is selected randomly as sensor nodes, i.e. to periodically transmit deadline-constrained data toward one of the k gateways. Each C_i is computed directly by the product of the time slot, i.e., 10 ms, and the number of hops in the path ϕ_i. T_i is harmonic and randomly generated in the range of $[2^4, 2^7]$, as in [4]. This implies a super-frame length of $H = 1280$ms. Finally, D_i is set implicitly, i.e. $D_i = Ti$.

*4) **Real-time assessment**:* We assess schedulability over a time interval equal to the super-frame, i.e., $\ell = H$, and when all the $m = 16$ channels are available. EDF and shortest path (Dijkstra) routing are assumed for all transmissions. Concerning $\Delta_{i,j}$, we use precise computation derived from the network topology.

B. Results & Discussion

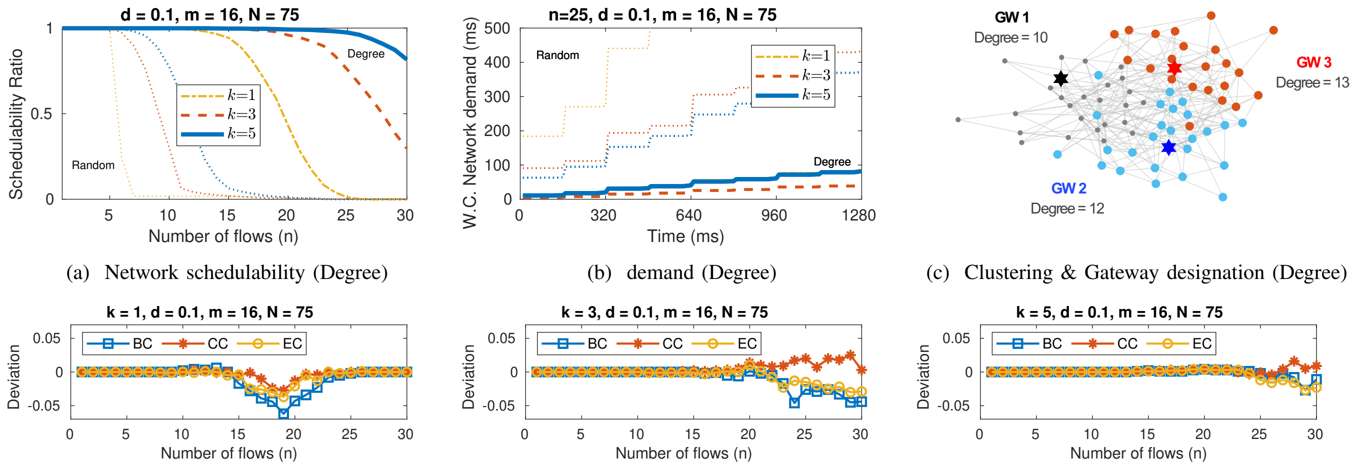
Fig. 1 shows simulation results based on the setup described above. Fig. 1a presents the schedulability ratio over n flows for both the proposed multi-gateway designation framework (thicker lines) and a random baseline (thinner lines). As expected, results show that increasing the number of gateways improves schedulability in all cases. A schedulability ratio of 99% can be achieved with only 5, 5 and 6 flows using random designation with k = 1, 3 and 5, respectively. With our framework, these values increase to 11, 17 and 21 flows, respectively, thus an improvement of 3.5 times in the number of schedulable flows with 5 gateways.

Fig. 1b shows the network demand for one of the 1000 random topologies imposed by $n = 25$ flows over an interval equal to the hyperperiod H, for both approaches. These results confirm that our framework significantly reduces the worst-case demand bounds.

Fig. 1c illustrates the clustering and gateway designation in a concrete topology when using our framework. The clustering is coded in different colours and the corresponding designated gateway is marked with a star.

Figures 1d, 1e and 1f present the deviation (difference) in the schedulability ratio achieved by the other centrality metrics w.r.t. degree centrality, namely, betweenness centrality (BC), closeness centrality (CC) and eigenvector centrality (EC), for $k \in \{1, 3, 5\}$. These results should be correlated with those in Fig. 1a. All centrality metrics perform equally well for low loads, given the low mutual interference, with all flows meeting their deadlines (100% schedulability). With

[1]*Notation*: $degree(v_q)$ denotes the number of edges of node v_q that are directly connected to any of the rest $N_\ell - 1$ nodes in G_ℓ; $sp_{r,s}$ is the number of shortest paths between any pair of cluster nodes v_r and v_s, and $sp_{r,s}(v_q)$ is the number of those paths passing through node v_q; $distance(v_p, v_q)$ is the (hop-count) shortest path distance between nodes v_p and v_q, with $p \neq q$, $\forall\ v_p \in V_\ell$, where V_ℓ is the set of vertices or nodes of cluster G_ℓ; x_j is the j-th value of the eigenvector x of the subgraph G_ℓ, and $\lambda_{max}(A_\ell)$ is the largest eigenvalue of the cluster's adjacency matrix $A_\ell = [a_{j,q}]_{N_\ell \times N_\ell}$, with $a_{j,q}$ being the matrix element at the row j and column q.



(a) Network schedulability (Degree) (b) demand (Degree) (c) Clustering & Gateway designation (Degree)

(d) $k = 1$ (e) $k = 3$ (f) $k = 5$

Fig. 1: a) schedulability ratio of 1000 random topologies with target density 0.1 and $k \in \{1, 3, 5\}$, with *degree* centrality versus random designation; b) worst-case network demand (ms) in one simulated case during a hyperperiod; c) illustrative example of joint clustering and gateway designation; d) e) and f) schedulability ratio deviation of other centrality metrics w.r.t. degree centrality for $k \in \{1, 3, 5\}$

.

high loads, mutual interference grows and all metrics perform poorly, with few or no flows meeting their deadlines (0% schedulability). The differences, in any case small ($< 5\%$), appear when schedulability starts degrading, thus being of low interest from a system design point of view.

VI. Conclusions

This paper presented a novel framework to design multi-gateway real-time WSN that improves traffic schedulability. We rely on *spectral clustering* unsupervised learning to provide k clusters without knowledge of nodes positions. In each cluster, the gateway is designated using *network centrality*. This generalizes the work in [4] for arbitrary gateways. Simulation results using random network topologies and flows with multiple configurations show the ability to reach 99% schedulability using 5 gateways with 3.5 times more flows when using our framework than using a random baseline. To the best of our knowledge, this is the first joint clustering and gateway designation method targeting schedulability in real-time TSCH WSNs. Future work will consider the use of the framework at run-time as well as the relationship between gateway designation, energy consumption and schedulability.

References

[1] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, “Industrial internet of things: Challenges, opportunities, and directions,” *IEEE Trans. on Industrial Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[2] D. Dujovne, T. Watteyne, X. Vilajosana, and P. Thubert, “6TiSCH: deterministic IP-enabled industrial internet (of things),” *IEEE Communications Magazine*, vol. 52, no. 12, pp. 36–41, 2014.

[3] C. Lu, A. Saifullah, B. Li, M. Sha, H. Gonzalez, D. Gunatilaka, C. Wu, L. Nie, and Y. Chen, “Real-time wireless sensor-actuator networks for industrial cyber-physical systems,” *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1013–1024, 2015.

[4] M. G. Gaitan, L. Almeida, A. Figueroa, and D. Dujovne, “Impact of network centrality on the gateway designation of real-time TSCH networks,” in *Proceedings of the 17th IEEE International Conference on Factory Communication Systems (WFCS)*, IEEE, 2021.

[5] F. Dobslaw, T. Zhang, and M. Gidlund, “QoS-aware cross-layer configuration for industrial wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1679–1691, 2016.

[6] M. C. Nascimento and A. C. De Carvalho, “Spectral methods for graph clustering–a survey,” *European Journal of Operational Research*, vol. 211, no. 2, pp. 221–231, 2011.

[7] Y.-S. Chen, S.-Y. Chang, T.-W. Chang, and M.-J. Tsai, “Multiple sink placement with latency and reliability guarantee in lossy wireless sensor networks,” in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, IEEE, 2018.

[8] M. Vahabi, H. R. Faragardi, and H. Fotouhi, “An analytical model for deploying mobile sinks in industrial internet of things,” in *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 155–160, IEEE, 2018.

[9] L. Mottola and G. P. Picco, “Muster: Adaptive energy-aware multisink routing in wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 10, no. 12, pp. 1694–1709, 2010.

[10] A. Lang, Y. Wang, C. Feng, E. Stai, and G. Hug, “Data aggregation point placement for smart meters in the smart grid,” *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 541–554, 2021.

[11] B. Xing, M. Deshpande, S. Mehrotra, and N. Venkatasubramanian, “Gateway designation for timely communications in instant mesh networks,” in *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 564–569, IEEE, 2010.

[12] M. G. Gaitan and P. M. Yomsi, “FF-DBF-WIN: On the forced-forward demand-bound function analysis for wireless industrial networks,” in *Proceedings of the Work-in-Progress Session of the 30th Euromicro Conference on Real-Time System (ECRTS)*, pp. 13–15, ECRTS, 2018.

[13] M. G. Gaitan, P. M. Yomsi, P. M. Santos, and L. Almeida, “Work-in-progress: Assessing supply/demand-bound based schedulability tests for wireless sensor-actuator networks,” in *16th IEEE Int. Conference on Factory Communication Systems (WFCS)*, pp. 1–4, IEEE, 2020.

[14] S. Baruah, V. Bonifaci, A. Marchetti-Spaccamela, and S. Stiller, “Improved multiprocessor global schedulability analysis,” *Real-Time Systems*, vol. 46, no. 1, pp. 3–24, 2010.

[15] C. Xia, X. Jin, and P. Zeng, “Resource analysis for wireless industrial networks,” in *Proceedings of the 12th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*, pp. 424–428, IEEE, 2016.

[16] A. Ng, M. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” *Advances in neural information processing systems*, vol. 14, 2001.

[17] T. W. Valente, K. Coronges, C. Lakon, and E. Costenbader, “How correlated are network centrality measures?,” *Connections*, vol. 28, no. 1, p. 16, 2008.