Optimal Routing Approaches for IEEE 802.15.4 TSCH Networks

Luigi Di Puglia Pugliese (A), Dimitrios Zorbas (B),
Francesca Guerriero (A), and Christos Douligeris (B)
(A) Department of Mechanical, Energy and Management Engineering, University of Calabria, Italy
(B) Department of Informatics, University of Piraeus, Greece
dzorbas@unipi.gr

(The current manuscript may not be the final version of the paper)

Abstract

The IEEE 802.15.4-Time Slotted Channel Hopping (TSCH) is a Medium Access Control (MAC) layer protocol designed for Industrial Internet of Things (IIoT) applications. TSCH focuses on the MAC layer only, while the construction of the routes relies on network layer protocols such as the Routing Protocol for Low-Power and Lossy Networks (RPL). The selection of the routes towards the sink plays a significant role in the nodes duty cycle, the delay, and the reliability of the network. In this paper, we formulate a multi-objective problem taking into account the schedule length (duty cycle/delay), the average number of hops (delay), and the cost of constructing the routes (reliability). To solve this problem, we use a scalarizing version where the objective function is defined as a convex combination of the three afore-mentioned parameters. Optimal computational as well as simulation results are presented. The findings of the current study can be used either as an optimal static routing solution when the link qualities are known and do not considerably change through time, or as a benchmark when designing low-power distributed protocols for TSCH networks.

1 Introduction

The advent of the fourth industrial revolution brings forward some new challenges in manufacturing and automation. Industrial Internet of Things (IIoT) applications use data generated by sensors, either built products out in the real world, or components within a machine on the production line, to impact how these things are manufactured, to minimize manufacturing mistakes and to reduce the production costs. Moreover, IIoT applications require efficient communications among the devices with low energy costs and high end-to-end reliability [25]. Time-slotted communications combined with a channel hopping MAC, such as the IEEE802.15.4 Time Slotted Channel Hopping (TSCH) [21] have been proposed to provide this required high reliability.

TSCH incorporates two major techniques that define the access to the medium; time synchronization and channel hopping. Time synchronization is used to achieve a very low-power operation. The nodes are synchronized according to a global clock initiated by the coordinator node (sink). The nodes wake up, sleep, and transmit at strict timings minimizing the waste of energy. Channel hopping is used to improve he reliability and cope with the external interferences. Successive transmissions are carried out at different frequencies increasing the probability of successfully delivering a frame that was not delivered in the previous transmissions.

In IEEE 802.15.4-TSCH networks, the time is divided in slotframes, each slotframe consisting of an equal number of timeslots. At any timeslot, a node may transmit a data packet, receive a data packet, or remain in sleep mode, thus saving significant amounts of energy. During the same timeslot, the sender of a data packet receives an acknowledgment packet if the transmission was successful. The timeslots that each pair of nodes remains active to transmit and receive data are defined by the scheduling. The scheduling is computed by either a centralized entity (e.g., the sink) or in a distributed manner by the nodes. The length of the schedule (i.e., the number of timeslots required to deliver all the data to the sink during a slotframe) affects the energy consumption of the nodes and the delay. The longer the schedule length, the higher the average delay and the energy consumption.

The Internet Engineering Task Force's (IETF) 6TiSCH Working Group, which is responsible for the standardization of the protocol stack for the IIoT, proposes the use of the IPv6 routing protocol RPL [24] in order to build and maintain the routes from the nodes to the sink. RPL is a distance-vector routing protocol specially designed by the ROLL IETF group for low-power devices and unreliable links. It setups a Destination Oriented Directed Acyclic Graph (DODAG) rooted at the sink (or at multiple sinks), and it assigns a rank to each node to define its virtual distance to the sink. An objective function is used to translate link metrics (e.g., hop count, latency) to a rank. Each node is, also, aware of a preferred parent through which it forwards its data to the next hop node towards the sink. However, RPL focuses on a distributed, low-cost construction and maintenance of the network topology and does not guarantee the optimality of the schedule. As we will see in the next section, an optimal schedule depends on some topology characteristics, such as the position of the nodes and how they are physically connected with each other. The default TSCH scheduling algorithm, called Scheduling Function Zero (SF0), as well as other recently proposed scheduling techniques [14, 8, 26, 4], are heavily dependent on the RPL mechanism in order to compute a data schedule. Nevertheless, as mentioned by Ioava et al. [11], a link quality metric that takes into account TSCH-specific statistics is still missing. Then scheduling statistics, such as the number of reserved cells, can be taken into account when defining the RPL rank of each node and appropriately select a default parent. It should be noted that the 6top sublayer [20], which is a 6TiSCH mechanism to dynamically build and adapt the schedule, allows neighboring nodes to add/delete TSCH cells to one another. This sublayer's operation can be also adjusted to balance the number of demanded cells so that neighboring parent nodes serve more or less equal number of packets.

From the theoretical point of view, it is important to mention that it is impossible to build a solution to minimize all the three network parameters simultaneously; the schedule length, the routing cost, and the number of hops. There are routing schemes that sometimes optimize both the schedule length and the packet routing cost [7] (e.g., the Minimum Spanning Trees). However, they tremendously increase the average hop count and, thus, the end-to-end delay in the network. As a consequence, the question to be answered is the following: *is there any routing solution with optimal schedule length which, at the same time, does not compromise (considerably) the packet routing cost and the hop count?*

We formulate an optimal routing tree problem as a minimization of an objective function that takes into account the schedule length, the average number of hops to the sink, and the packet routing cost. We solve the problem by using a scalarizing version where the objective function is defined as a convex combination of the three parameters. We compare the results obtained with the model to the standard RPL protocol using different node and link evaluation metrics, i.e., the hop count, the expected transmission count (ETX), and the number of children (i.e., the predecessor nodes of a node towards the sink). We also compare the performance of the proposed approach to that of a non-distance vector protocol which uses a Connected Dominating Set (CDS) to compute the core of the network while the rest of the nodes are 1-hop cluster members of the CDS nodes.

This is the first paper that studies the problem of routing with optimal schedules in TSCH networks. Its objective is to shed light – from the theoretical perspective – on how the nodes could be organized in order to achieve optimal schedule lengths while not burdening considerably or even reducing the overall packet routing cost and the average hop count. In particular, we show that the current RPL metrics perform far from the optimal in terms of schedule length; a behavior that can be improved if the number of children is included in the parent selection process. Moreover, the theoretical findings reveal that a balanced tree with less, on the average, children per parent exhibit far better results in terms of schedule length than the standard RPL solutions. Finally, the the cluster-based approach seems not to be appropriate for use in TSCH networks since all the evaluated parameters worsen considerably in this approach.

The rest of this paper is organized as follows. Section 2 describes the scheduling process in IEEE802.15.4-TSCH networks and how the schedule length is affected by the applied routing solution. In Section 3, we formulate an optimal routing tree problem as a function of the schedule length, the cost, and the hop count. Computational as well as simulation results are presented and discussed in Section 4. Finally, Section 5 concludes the paper and presents ideas for future enhancements.

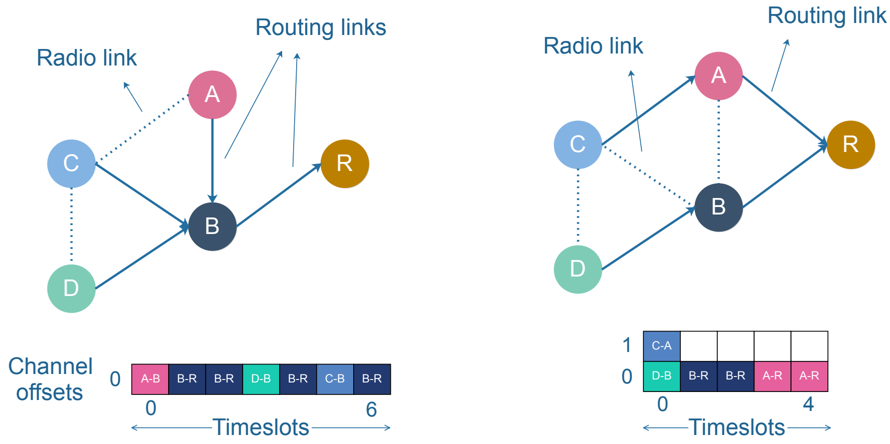
2 The IEEE802.15.4-TSCH protocol

2.1 Data transmission scheduling

Scheduling is one of the fundamental operations in IEEE802.15.4-TSCH networks. The basic operation of a scheduler is to define when (i.e., at which timeslot) a node must turn on its radio to transmit or receive a data packet. To do this, a scheduler has to consider the number of packets queuing at each node in each timeslot and the fact that a node cannot transmit and receive a data packet in the same timeslot. Moreover, the scheduler assigns a channel offset to each node for each of the participating links. This offset is then translated to a physical frequency according to a function described in the standard [21]. Physical neighboring links as well as 2-hop neighbors on the same path to the sink must have a different channel offset to avoid collisions when they transmit in the same timeslot. The job of a scheduler – centralized or distributed – is to take into account these constraints and compute a collision-free schedule with the shortest possible length. The length of the schedule is measured in number of timeslots.

2.2 How routing affects scheduling

The topology characteristics and the way the nodes are connected with each other may affect the scheduling length. The optimal length is lower-bounded by the number of packets generated in the network during one



(a) A possible routing tree (b) An optimal routing tree

Figure 1: A possible and an optimal routing solution in terms of schedule length for a network with 5 nodes.

slotframe since only one packet is sent per timeslot. This requires that the sink receives one packet per timeslot during the entire slotframe. However, due to the bottleneck effect of the 1-hop neighbors of the sink, the latter may not receive a packet in one or more timeslots.

To better understand this issue, we introduce the example of Figure 1, where we display two different routing solutions for a 5-node network. The corresponding data transmission schedules are illustrated below each solution. We assume that all the links present a similar quality (reliability) and all the nodes generate 1 packet per slotframe which must be forwarded to the sink via an 1-hop or multi-hop manner. This results in a total number of 4 generated packets while the total number of transmissions may vary depending on the routing approach. The bold lines correspond to the routing links and the dash lines to the radio links. The radio links correspond to the physical connections between the nodes, while the routing links are physical connections that can be used exclusively for data transmissions. On one hand, the solution on the left exhibits a schedule length equal to 7 slots since 7 transmissions need to be performed in order to forward all the packets to the sink. Three out of the seven slots are reserved for transmissions from nodes A, C, D to B (i.e., timeslots 0, 3, and 5). Since in TSCH a node can only receive or transmit a packet during a timeslot, R (i.e., the sink) is inactive in these three timeslots. On the other hand, the routing tree on the figure on the right exhibits an optimal performance with 5 timeslots reserved in total. This solution presents only one inactive timeslot for the sink (i.e., timeslot 0) since the transmissions of C and D can happen in parallel using different channel offsets.

If we now assume that links C-B and A-B present better quality characteristics than C-A and A-R, the solution on the left exhibits in total a better tree cost compared to the solution on the right. In this case, it is straightforward to observe that it is unlikely to find a tree that optimizes both the cost and the schedule length. Generalizing, if at least one pair of links presents the above behavior, the minimization of the one measurement does not guarantee the minimization of the other and vice-versa.

3 Problem formulation

In this section we mathematically formulate the packet routing problem in IEEE802.15.4-TSCH networks. In order to mathematically study the routing behavior, we assume stable environmental conditions. This implies that no parent and routing changes are allowed once the routes have been initially computed. Therefore, RPL route maintenance capabilities, like the trickle algorithm [12] or the hysteresis function [10], are not evaluated in this paper.

The problem is defined over a graph $G(N, E)$, where N is the set of nodes and E is the set of edges. Each edge $[i, j]$ models the radio link between node i and j. A cost l_{ij} is associated with each edge $[i, j] \in E$ as given by Eq.(1):

$$l_{ij} = \frac{dist(i,j)}{max\_dist}, \tag{1}$$

where $dist()$ is the Euclidean distance function between two nodes and max_dist is the least reliable (maximum) distance between two nodes. This distance can be approximated experimentally. Since link reliability generally worsens with distance [15], in the present paper, we define the packet routing cost as a function of the relative distance between the nodes that constitute the edges of the data collection tree and the number of packets routed by each node. This means that the longer the distance and the higher the number of packets, the higher the routing cost. Thus, the cost of Eq. (1), whose value increases with the distance, is used as a simple link evaluation measurement, expressing the tolerance of the link to packet losses. This cost function is used only for modeling purposes and not as a real indicator to assess the packet reception probability.

The number of packets q_i to be sent to the *sink* node is associated with each node $i \in N \setminus \{sink\}$. The aim of the formulation is to determine the routing of the packets $q_i, \forall i \in N \setminus \{sink\}$, where the data, starting from node i, reach node *sink* possibly passing through other nodes of the graph.

The problem can be viewed as a particular instance of the multi-commodity flow problem. In general, the problem is formulated on a directed graph and a source node, a destination node, and the quantity of items to be sent are associated with each commodity. The aim is to send the quantity of items associated with each commodity from the associated source node to the associated destination node, by minimizing the overall transportation cost and guaranteeing the satisfaction of a resource constraint associated with each arc. Indeed, the arcs are capacitated, thus, the quantity of items passing through each arc must be less than or equal to a given threshold.

In order to formulate the problem at hand as an instance of the multi-commodity flow problem, we construct a directed graph $D(N, A)$ where A is the set of arcs. In particular, we associate with each edge $[i, j] \in E$ two arcs, i.e., arcs (i, j) and (j, i) whose cost is l_{ij}, which are stored in the set A. The problem has $|N| - 1$ commodities. Each commodity k is associated with a node $i \in N \setminus \{sink\}$. Thus, it is characterized by the quantity of items q_i, the source node i and the destination node *sink*. Thus, the flow associated with each commodity is the number of packets that traverses the graph from node i to node *sink*. It is worth observing that an arc $(i, j) \in A$ is possibly shared by the commodities. This means that the flows associated with several commodities can transit on the some arcs. However, no limitation is given on the number of packets that can transit through each arc.

An edge $[i, j] \in E$ is active if at least one commodity uses an arc associated with it. The set of active edges represents a tree T of G. The notations and the parameters used to model the problem are listed below:

- $G(N, A)$: directed graph;
- N: set of nodes;
- q_i: quantity of packets to be sent from node i to node *sink* through graph G.
- A: set of arcs;
- l_{ij}: cost associated with arc $(i, j) \in A$.

Before introducing the objective functions along with the mathematical formulation, we introduce the following decision variables:

- $y^k_{ij}, \forall (i, j) \in A, k \in N \setminus \{sink\}$: binary variables stating whether the flow of commodity k, that is, the packets associated with the node k, transits through arc $(i, j) \in A$.
- $x_{ij}, \forall (i, j) \in A$: binary variables indicating whether arc (i, j) is active, that is, $y^k_{ij} = 1$ for at least one commodity k.
- $Q_i = \sum_{(u,v)\in T_i} q_u + q_i \; \forall i \in N$: continuous variables representing the number of packets that reach node $i \in N$, where T_i is the subtree of T rooted at node i.

The cost of a tree T is defined as the sum of the cost multiplied by the packets associated with the active edges, i.e., $c(T) = \sum_{(i,j)\in T} l_{ij} Q_i$. In this way, we evaluate each link based on the amount of traffic passing through this particular link.

The schedule length LT can be modeled as follows [13]:

$$LT = \max\{2Q_M - q_M, Q_{sink}\}, \tag{2}$$

where $M = \arg\max_{i\in N}\{Q_i : x_{isink} = 1\}$. In other words, M is the node, among all those directly connected to the *sink*, associated with the maximum value of variable Q. In order to properly take into account function LT, we define a binary variable z_i, $\forall i \in N$ indicating whether M, in Eq. (2), is defined for node $i \in N$.

We consider also the average number of hops as a parameter to take into account in the determination of the optimal solution. In particular, we compute the average number of hops H as $\sum_{k\in N\setminus\{sink\}} \sum_{(i,j)\in A} \frac{1}{|N|-1} y^k_{ij}$.

Based on the above we get the following model:

$$\min C = \sum_{(i,j)\in A} l_{ij} Q_i x_{ij} \tag{3}$$

$$\min LT \tag{4}$$

$$\min H = \sum_{k\in N\setminus\{sink\}} \sum_{(i,j)\in A} \frac{1}{|N|-1} y_{ij}^k \tag{5}$$

$s.t.$

$$\sum_{j:(i,j)\in A} y_{ij}^k - \sum_{j:(j,i)\in A} y_{ji}^k = \begin{cases} 1 & \text{if } i = k \\ -1 & \text{if } i = sink \quad \forall k \in N \setminus \{sink\}, \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$x_{ij} \geq y_{ij}^k, \quad \forall (i,j) \in A, k \in N \setminus \{sink\}, \tag{7}$$

$$LT \geq Q_{sink}, \tag{8}$$

$$LT \geq 2Q_M - \sum_{i\in N:(i,sink)\in A} z_i q_i, \tag{9}$$

$$Q_i \geq \sum_{(v,i)\in A} Q_v x_{vi} + q_i \quad \forall i \in N, \tag{10}$$

$$Q_M \geq Q_i - Q(1 - x_{isink}), \quad \forall i \in N, \tag{11}$$

$$z_i Q_M \leq Q_i, \quad \forall i \in N : (i, sink) \in A, \tag{12}$$

$$\sum_{i\in N:(i,sink)\in A} z_i = 1, \tag{13}$$

The objective function (3) minimizes the total packet routing cost, (4) minimizes LT and (5) optimizes the average number of hops. Eq. (6) define the flow conservation constraints for each commodity k. Eq. (7) define the value of variables x. Eq. (8) and (9) linearize function LT. Eq. (10) define the value of Q_i. Eq. (11) set Q_M to the maximum Q_i among all nodes i directly connected to the $sink$. Eq. (12) and (13) define the value of z_i. In particular, if $Q_i \neq Q_M$, then $z_i = 0$. On the other hand, constraints (12) and (13) are satisfied only for $z_i = 1$ with $Q_i = Q_M$. The parameter Q is a large number that can be set equal to $\sum_{i\in N} q_i$.

The Constraints (10) and (12) as well as the objective function (3) are not linear due to the terms $Q_v x_{vj}$ and $z_i Q_M$, respectively. The former can be linearized by introducing non-negative auxiliary variables $\hat{Q}_{vj}$. The linearized objective function (3) and the constraints (10) take the following form:

$$\sum_{(i,j)\in A} l_{ij} \hat{Q}_{ij}, \tag{14}$$

$$Q_i \geq \sum_{(v,i)\in A} \hat{Q}_{vi} + q_i \quad \forall i \in N, \tag{15}$$

respectively. The variables $\hat{Q}_{ij}$ are defined such that if $x_{ij} = 1$, then $\hat{Q}_{ij} = Q_i$. Otherwise, if $x_{ij} = 0$, $\hat{Q}_{ij} = 0$. To take into account the aforementioned definition, the following constraints are added to the model

$$\hat{Q}_{ij} \geq Q_i - Q(1 - x_{ij}), \quad \forall (i,j) \in A, \tag{16}$$

$$\hat{Q}_{ij} \leq Q x_{ij}, \quad \forall (i,j) \in A, \tag{17}$$

$$\hat{Q}_{ij} \leq Q_i, \quad \forall (i,j) \in A. \tag{18}$$

Constraints (11) can be simplified in the following way:

$$Q_M \geq \hat{Q}_{isink}, \quad \forall (i, sink) \in A, \tag{19}$$

To linearize the terms $z_i Q_M$, we introduce the non-negative variables $z_i^M, \forall i \in N : (i, sink) \in A$, which are defined as follows:

$$z_i^M \geq Q_M - Q(1-z_i), \;\; \forall i \in N : (i, sink) \in A, \tag{20}$$
$$z_i^M \leq Q z_i, \;\; \forall i \in N : (i, sink) \in A, \tag{21}$$
$$z_i^M \leq Q_M, \;\; \forall i \in N : (i, sink) \in A. \tag{22}$$

The constraints (12) are then formulated as follows:

$$z_i^M \leq Q_i, \;\; \forall i \in N : (i, sink) \in A. \tag{23}$$

The proposed multi-objective formulation does not provide a single optimal solution,but more than one Pareto optimal solutions exist. Indeed, as mentioned before, the minimization of the cost does not guarantee the minimization of both LT and average number of hops H and vice-versa.

In general, the number of Pareto optimal solutions grows exponentially with the size of the problem [3, 2, 6, 17]. Generating all Pareto optimal solutions becomes intractable for high dimension instances. In addition, the decision maker has to choose, among all Pareto optimal solutions, the one that satisfies the requirements. However, the higher the number of possible solutions, the more difficult the choice of the most satisfactory one.

To overcome these drawbacks, scalarizing techniques can be used. The main idea is to combine the multiple objectives into one single-objective scalar function and to solve the multi-objective problem as a single-objective optimization problem. Many different scalarizing functions have been proposed in the scientific literature, i.e., the Chebyshev norm, the utility functions, and the weighted sum method [22, 18].

In the first case, a reference point, representing the preferences of the decision-maker in the criteria space, is considered. Optimizing this function means to obtain a Pareto optimal solution that is closest to the reference point with respect to the Chebyshev norm. The decision maker can, thus, have a better overview of the objective function values.

In the utility function techniques, a utility function is constructed, on the basis of the users' preferences, and an optimization problem is solved to find the solution that maximizes this function.

We use the weighted sum method to address the problem studied in this paper. The main aim is to build an aggregate objective function, obtained by summing up all the objective functions, each of them multiplied by a weighting factor. In particular, we consider a convex combination of the three parameters, i.e., C, LT and H, as follows:

$$\min \alpha \sum_{(i,j)\in A} l_{ij}\hat{Q}_{ij} + \beta LT + \gamma \sum_{k \in N\setminus\{sink\}} \sum_{(i,j)\in A} \frac{1}{|N|-1} y_{ij}^k \tag{24}$$

where $\alpha, \beta, \gamma \in [0, 1]$ and $\alpha + \beta + \gamma = 1$. It is worth observing that, in general, the Pareto front is not convex, thus, the Pareto solutions that do not belong to the convex hull cannot be computed when minimizing function (24).

Our choice of the weights has been motivated by the fact that this strategy represents the most widely-used method in the multi-objective optimization [9] and it can handle problem with any number of objective functions.

The optimal solution is a tree T, that is, a connected and acyclic graph. Thus, there exists a unique path from each node $k \in N$ to the $sink$. Varying the parameters α, β, γ, different trees, belonging to the convex hull, can be obtained.

Since C, LT, and H are of different nature, we have to consider the scaling coefficients s_c, s_{lt}, and s_h in function (24). Let S^c, S^{lt}, and S^h be the optimal solution to the problem when optimizing function (3), (4), and (5), respectively. The scaling coefficients are computed as $s_c = 1/\left(\max\{C(S^{lt}), C(S^h)\} - C(S^c)\right)$, $s_{lt} = 1/\left(\max\{LT(S^c), LT(S^h)\} - LT(S^{lt})\right)$, and $s_h = 1/\left(\max\{H(S^c), H(S^{lt})\} - H(S^h)\right)$ for function C, LT, and H, respectively.

4 Evaluation & discussion of the results

In this section we compare the results derived by solving the proposed model with the results of two well-known routing approaches. The first approach is the distance vector routing protocol RPL and the second one is a clustering approach based on the construction of a CDS [1]. We distinguish three metrics to construct the routes for RPL and CDS; the number of hops to the sink, the ETX [5], and the number of children of the parent nodes. These are link and node quality metrics proposed by IETF [19]. For the last metric the objective function proposed in [16] is used in order to distribute the load among a higher number of parents. These three metrics are used by the objective function of RPL and are translated to a rank. In CDS, they are used as edge weights to compute the connections between the dominating nodes (cluster heads). They are also used as a metric for the cluster nodes to connect to the cluster head with the minimum cost.

ETX is computed by taking into account 10 successive packet transmissions. We consider a transmission as successful if the signal power P_{tx} at the destination is higher than the required sensitivity power, which is fixed by the RF manufacturer, as it is described by Eq. (25) [23]:

$$P_{tx}\frac{e^{2\sigma G}}{d^{2b}} > S, \tag{25}$$

where P_{tx} is the transmission power, S is the sensitivity power, and $e^{2\sigma G}$ has a log-normal distribution with a shadowing coefficient σ ($G \sim N(0,1)$). The term $1/d^{2b}$ accounts for the far-field path loss with distance d, where the amplitude loss exponent b is environment-dependent. We must note here that MAC layer phenomena, such as packet time outs or CCA failures are not taken into account.

4.1 Setup

We consider scenarios with different node populations, ranging from 20 to 100, and we create 25 instances with random node positions for each scenario. We use a terrain size of 100×100 meters and we guarantee that all the nodes can reach the sink in an 1-hop or in a multi-hop manner. We assume that there is a link between any two nodes if the signal power at the receiver is higher than the required sensitivity power. Each node generates from 1 to 3 packets per slotframe. Regarding the node parameters we rely on the characteristics of Zolertia RE-motes[1] with a typical transmission power of 0dBm. The maximum communication range varies according to the sensitivity of the antenna and the propagation model. We use the following values which correspond to outdoor communications: σ=0.7 ($\sigma_{dB} = 6dB$), random $G \in (0,1)$, and b=2 (exponent of 4) [23]. We implement RPL and CDS in the Perl programming.

With regards to the model, we conduct the experimental phase considering three different values for each of α, β, and γ, that is $\{0.2, 0.3, 0.5\}$. Thus, we consider six combinations of those parameters resulting in 150 instances for each scenario. We implement the model in the Java programming language and the solver CPLEX 12.51 is called for its resolution. The tests are carried out on an Intel(R) Core(TM) i7-4720HQ CPU with 8GB of RAM under the Microsoft Windows 10 operating system. From a wide range of preliminary computational results, we obtained that CPLEX is not able to compute an optimal solution for the instances with a number of nodes greater than or equal to 30 in a reasonable amount of time. For this reason, we focus our attention on the instances with 20, 30, and 40 nodes imposing a time limit of 30 minutes. In the case the solver does not find an optimal solution within the imposed time limit, then the best feasible solution is returned. According to our experiments, the chosen time limit represents a good compromise between the quality of the obtained solution and the computational overhead.

4.2 Results

The average numerical results along with the simulation results over 25 instances are reported in Table 1. The value of the cost per unit of nodes is reported under column $\frac{C}{|N|-1}$, the schedule length is reported under column LT, and the average number of hops is reported under column H. Table 1 shows the aforementioned results for each considered combination of α, β, and γ for the model, and for each metric, that is, for the hops, ETX, and the children for RPL and CDS. The rows min, max, and AVG report the minimum, the maximum, and the average values for $\frac{C}{|N|-1}$, LT, and H. Figures 2 and 3 present the individual RPL and CDS results respectively with different link metrics. A comparison with the best values derived from the model (for 20, 30, and 40 nodes) is also displayed.

We must note that the solver returns 114 optimal solutions within the imposed time limit for the scenario with 20 nodes. Whereas, only feasible solutions are available for the scenario with 30 and 40 nodes.

The results collected by the model highlight that the obtained solution is not highly affected by the parameters α, β, and γ for the considered instances. The maximum values of $\frac{C}{|N|-1}$, LT, and H, averaged over all instances and value of α, β, and γ, are 3.2%, 2.9%, and 3.4% higher than the minimum values, respectively. However, we have to notice that this gap increases when the number of nodes increases. The highest gap is obtained for $\frac{C}{|N|-1}$ and for the instances with 30 nodes, with a value of 5.3%. For LT and H, the maximum gap is 4% and 5%, respectively, for the instances with 40 nodes.

As expected, the higher the number of nodes, the higher the value of both $\frac{C}{|N|-1}$ and LT (see rows AVG of Table 1). This trend is not affected by the value of H.

RPL and CDS perform significantly worse than the model in terms of schedule length. Indeed, the values of LT, considering the metrics "hops", "ETX", and "children" obtained with RPL, are 1.10, 1.16, and 1.08 times

[1] https://github.com/Zolertia/Resources/wiki/RE-Mote

nodes				model			metrics	RPL			CDS-based		
	α	β	γ	$\frac{C}{\lvert N\rvert-1}$	LT	H		$\frac{C}{\lvert N\rvert-1}$	LT	H	$\frac{C}{\lvert N\rvert-1}$	LT	H
20	0.2	0.3	0.5	4.732	46.720	3.198	hops	4.760	49.160	3.020	4.793	51.000	3.402
20	0.2	0.5	0.3	4.732	46.640	3.198							
20	0.3	0.2	0.5	4.732	46.720	3.198	ETX	4.844	53.240	3.245	4.898	54.400	3.718
20	0.3	0.5	0.2	4.740	46.680	3.204							
20	0.5	0.2	0.3	4.726	46.680	3.200	children	4.779	50.160	3.020	4.854	51.120	3.476
20	0.5	0.3	0.2	4.729	46.640	3.200							
min				4.726	46.640	3.198		4.760	49.160	3.020	4.793	51.000	3.402
max				4.740	46.720	3.204		4.844	53.240	3.245	4.898	54.400	3.718
AVG				4.732	46.680	3.200		4.794	50.853	3.095	4.848	52.173	3.532
30	0.2	0.3	0.5	4.928	65.520	3.385	hops	4.855	75.880	3.105	5.078	75.840	3.560
30	0.2	0.5	0.3	4.830	64.120	3.337							
30	0.3	0.2	0.5	4.686	66.160	3.243	ETX	4.856	79.320	3.490	5.466	84.480	4.300
30	0.3	0.5	0.2	4.700	64.640	3.277							
30	0.5	0.2	0.3	4.703	65.920	3.279	children	4.809	73.440	3.105	5.522	78.960	3.900
30	0.5	0.3	0.2	4.680	65.320	3.252							
min				4.680	64.120	3.243		4.809	73.440	3.105	5.078	75.840	3.560
max				4.928	66.160	3.385		4.856	79.320	3.490	5.522	84.480	4.300
AVG				4.755	65.280	3.296		4.840	76.213	3.233	5.355	79.760	3.920
40	0.2	0.3	0.5	4.691	93.400	3.190	hops	4.936	99.880	3.029	5.488	103.840	3.713
40	0.2	0.5	0.3	4.869	89.800	3.345							
40	0.3	0.2	0.5	4.681	91.520	3.199	ETX	4.931	104.480	3.424	5.879	113.080	4.458
40	0.3	0.5	0.2	4.681	93.400	3.204							
40	0.5	0.2	0.3	4.690	93.160	3.187	children	4.876	97.520	3.029	5.916	110.800	3.964
40	0.5	0.3	0.2	4.722	91.160	3.258							
min				4.681	89.800	3.187		4.876	97.520	3.029	5.488	103.840	3.713
max				4.869	93.400	3.345		4.936	104.480	3.424	5.916	113.080	4.458
AVG				4.722	92.073	3.231		4.914	100.627	3.161	5.761	109.240	4.045

Table 1: Numerical and simulation results for the model with a linear combination of C, LT and H compared to RPL and the CDS-based approach with different link metrics.

higher than the average value of LT returned by the model, respectively. RPL achieves slightly worse values of $\frac{C}{N-1}$ than the model for all metrics, on average. Referring to the values of H, RPL behaves slightly better than the model when considering both the "hops" and the "children" metrics.

CDS exhibits the worst behaviour. The value of LT is 1.13, 1.23, and 1.18 times higher than that obtained by the model, considering the metrics "hops", "ETX", and "children", respectively. This trend is observed for the values of both LT and H.

Overall, RPL behaves better than CDS for the instances with 20, 30 and 40 nodes. This can be also observed from the simulation results with higher node populations presented in Figures 2 and 3 for the RPL and the CDS cases.

The difference between the two approaches is higher when more nodes are added in the terrain. Indeed, CDS needs up to 11% more timeslots and uses up to 23% longer routes in terms of number of hops compared to RPL. However, CDS uses on average slightly shorter links which might result in more reliable communications.

Comparing the link metrics in RPL, as it was expected, ETX produces shorter links at the cost of longer schedules and slightly higher hop counts. When the number of children is used as a metric to select the default parent, RPL produces the best results among the three metrics. This happens because this metric tends to balance the number of nodes between the parents, thus decreasing the appearance of bottlenecks.

Figure 4 illustrates the solutions of the three approaches for a scenario with 30 nodes. The optimal solution in terms of schedule length is 57 timeslots achieved by all the examined α, β, γ combinations of the model. RPL achieved 86 and 79 timeslots with ETX and "children", respectively. The cluster approach achieved 111 timeslots with ETX and 95 with "children". The figure reveals that the optimal solution exhibits a more balanced tree with less bottlenecks compared to the other two approaches.

Considering the rest of the instances as well, the model generates solutions with a higher, on the average, number of 1-hop to the sink nodes, while the Q values of these neighbors are comparable with each other. This is practically translated to less inactive timeslots for the sink, which is the cause of the schedule length growth. We summarize these results in Table 2 where the number of sink neighbors is reported under column NB and the variance associated with the value of Q is reported under column σ. We observe that the model produces

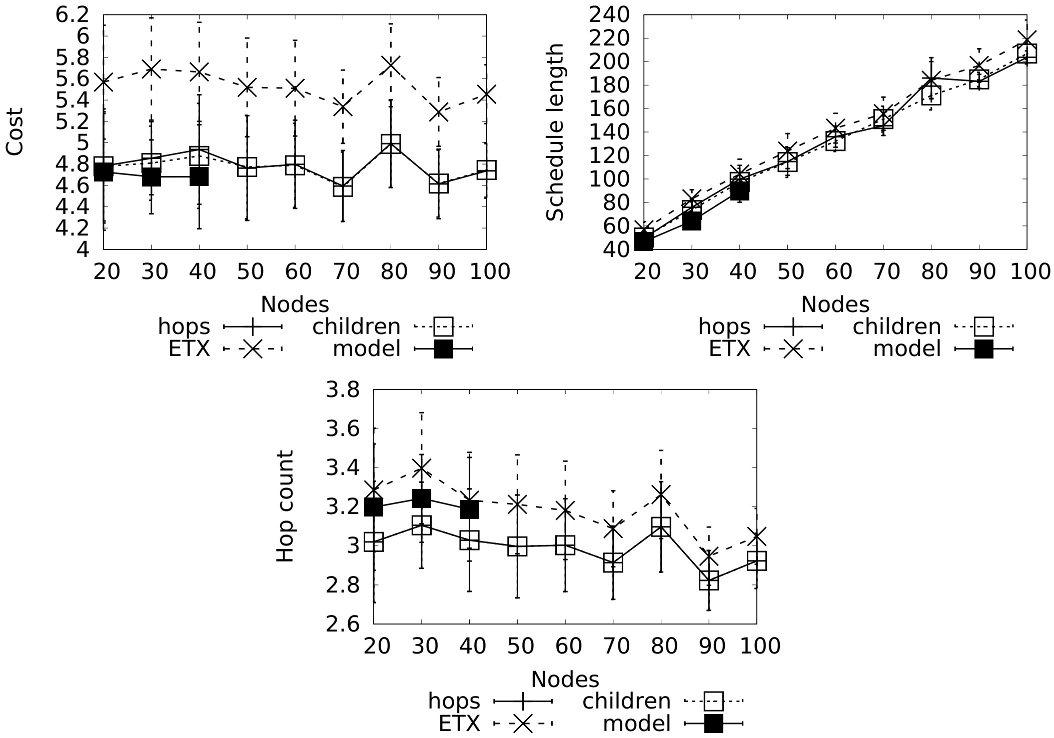


Figure 2: RPL performance with different link evaluation metrics and variable number of nodes.

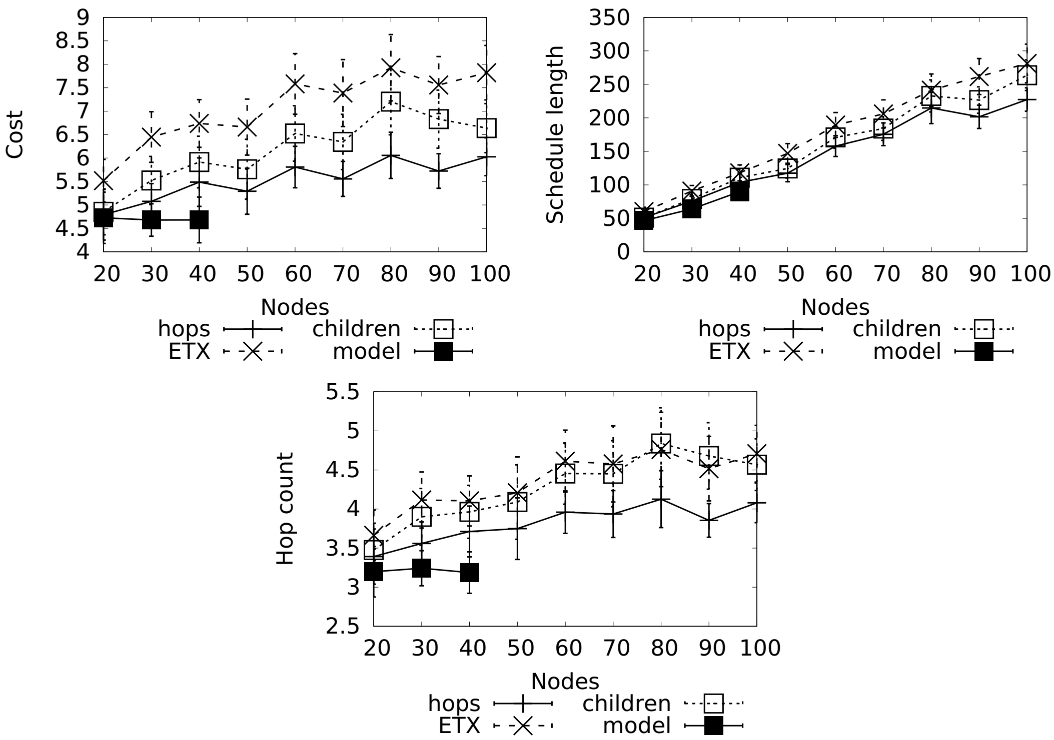
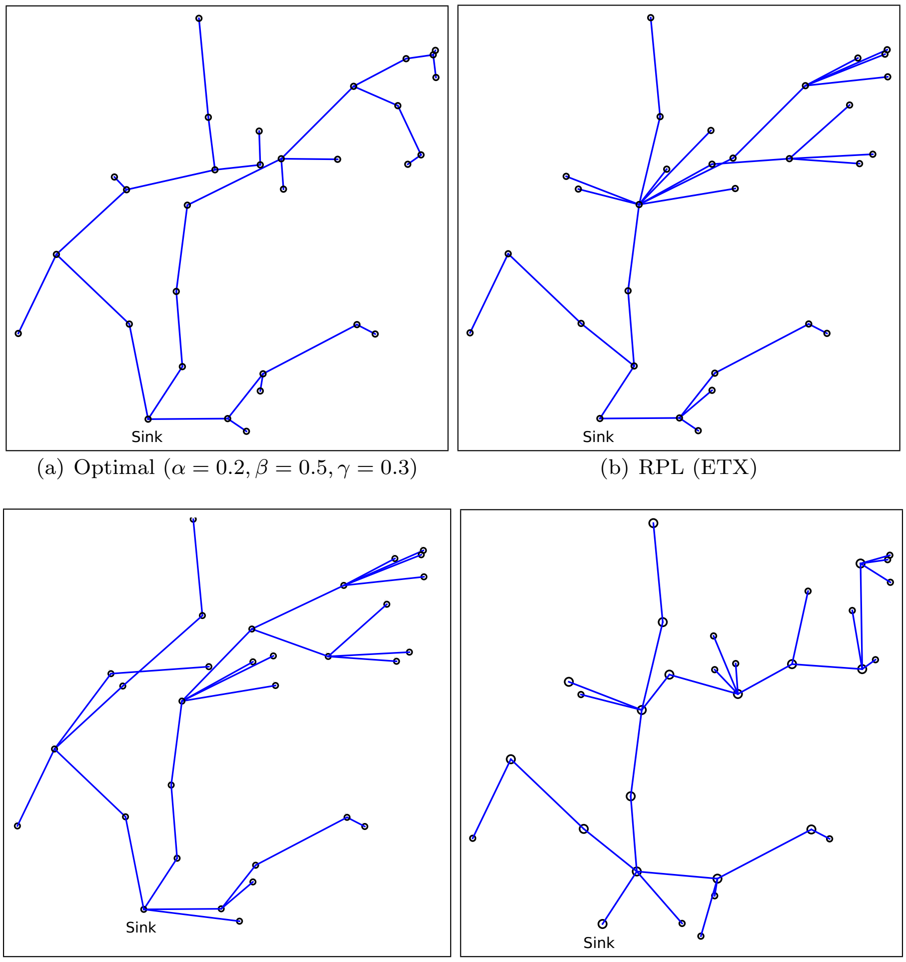


Figure 3: CDS-based routing performance with different link evaluation metrics and variable number of nodes.



(a) Optimal ($\alpha = 0.2, \beta = 0.5, \gamma = 0.3$) (b) RPL (ETX) (c) RPL (children) (d) CDS-based (ETX)

Figure 4: The RPL and CDS-based routing trees compared to the optimal solution for a 30 nodes instance .

results with a higher number of sink neighbors (NB) and their Q values present a lower variance (σ). The "hops" and the "children" metrics exhibit NB values similar to the model, however, the variance has a range from 12 to 29% higher. The ETX metric exhibits a worse performance since many nodes in a neighborhood tend to stick with the same parent overloading specific tree branches. This behavior has been also captured in [16].

5 Conclusions & Future work

The present paper examined the problem of optimal routing in TSCH networks. We explained the connection between the routing process and the schedule length and we showed that an optimum routing tree in terms of schedule length may not be optimal in terms of cost and vice-versa. We defined a multi-objective optimization problem as a function of the schedule length, the packet routing tree cost, as well as the hop count, a factor that heavily affects the end-to-end delay of the network. We solved the problem by considering a scalarizing version where the objective function is defined as a convex combination of the three parameters. The comparison with the RPL protocol using three wide-used link/node quality metrics as well as with a CDS-based approach reveal that there is a great room for improvement in terms of schedule length without compromising the tree cost and the average hop count.

The findings encourage the enhancement of the existing RPL metrics with information regarding the number of children per parent in order to construct more balanced routing trees. However, the optimal number of children per parent is difficult to be computed in a distributed manner since it requires knowledge of some global parameters, such as of the network density. A possible solution would be to gradually adjust the routing tree by including scheduling statistics from the 6top layer into the RPL broadcast messages. Finally, computing optimal routing solutions in the presence of multiple sinks is an open problem. In the future, we will investigate how to organize the nodes in multiple trees such as each of these trees exhibits an optimized schedule length as well as low cost and latency.

Nodes	model					metric	RPL	
	α	β	γ	NB	σ		NB	σ
20	0.2	0.3	0.5	4.08	6.638	hops	4.080	7.816
20	0.2	0.5	0.3	4.08	6.642			
20	0.3	0.2	0.5	4.08	6.651	ETX	3.440	8.266
20	0.3	0.5	0.2	4.04	6.453			
20	0.5	0.2	0.3	4.08	6.668	children	4.080	7.449
20	0.5	0.3	0.2	4.08	6.668			
30	0.2	0.3	0.5	5.04	10.803	hops	5.040	13.200
30	0.2	0.5	0.3	5.04	10.615			
30	0.3	0.2	0.5	5	11.060	ETX	3.880	14.234
30	0.3	0.5	0.2	5	10.885			
30	0.5	0.2	0.3	5.04	10.978	children	5.040	13.564
30	0.5	0.3	0.2	5.04	10.815			
40	0.2	0.3	0.5	6.6	13.687	hops	6.760	14.961
40	0.2	0.5	0.3	6.6	12.674			
40	0.3	0.2	0.5	6.56	13.423	ETX	5.120	15.492
40	0.3	0.5	0.2	6.64	13.572			
40	0.5	0.2	0.3	6.64	13.476	children	6.760	15.032
40	0.5	0.3	0.2	6.6	12.994			

Table 2: The number of sink neighbors (NB) and the standard deviation (σ) of their Q values.

Acknowledgments

This work was carried out within the action "Strengthening Post Doctoral Research" of the "Human Resources Development Program, Education and Lifelong Learning", 2014-2020, which is being implemented from IKY and is co-financed by the European Social Fund – ESF and the Greek government.

References

[1] J. Carle and D. Simplot-Ryl. Energy-efficient area monitoring for sensor networks. *Computer*, 37(2):40–46, Feb 2004.

[2] Z. Chinchuluun and P. M. Pardalos. A survey of recent developments in multiobjective optimization. *Ann Oper Res*, 154:29–50, 2007.

[3] J. R. Current and H. Min. Multiobjective network design of transportation networks: Taxonomy and annotation. *European Journal of Operational Research*, 26:187–201, 1986.

[4] Glenn Daneels, Bart Spinnewyn, Steven Latré, and Jeroen Famaey. Resf: Recurrent low-latency scheduling in ieee 802.15. 4e tsch networks. *Ad Hoc Networks*, 69:100–114, 2018.

[5] Douglas S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris. A high-throughput path metric for multi-hop wireless routing. In *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (Mobicom), pages 134–146, San Diego, USA, 2003. ACM.

[6] L. Di Puglia Pugliese and F. Guerriero. Dynamic programming approaches to solve the shortest path problem with forbidden paths. *Optim Method Softw*, 28(2):221–255, 2013.

[7] L. Di Puglia Pugliese, D. Zorbas, and F. Guerriero. Modeling and solving the packet routing problem in industrial iot networks. In *2018 International Conference on Optimization and Decision Science*, 2018 International Conference on Optimization and Decision Science, Tormina, Italy, Sep 2018. Springer.

[8] Simon Duquennoy, Beshr Al Nahas, Olaf Landsiedel, and Thomas Watteyne. Orchestra: Robust mesh networks through autonomously scheduled tsch. In *Proceedings of the 13th ACM conference on embedded networked sensor systems*, Proceedings of the 13th ACM conference on embedded networked sensor systems, pages 337–350. ACM, 2015.

[9] Matthias Ehrgott and Xavier Gandibleux. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR-Spektrum*, 22(4):425–460, 2000.

[10] O. Gnawali. The minimum rank with hysteresis objective function. RFC 6719, 2012.

[11] O. Iova, F. Theoleyre, T. Watteyne, and T. Noel. The love-hate relationship between ieee 802.15.4 and rpl. *IEEE Communications Magazine*, 55(1):188–194, January 2017.

[12] P. Levis and T.H. Clausen. The trickle algorithm. RFC 6206, 2011.

[13] M. R. Palattella, N. Accettura, L. A. Grieco, G. Boggia, M. Dohler, and T. Engel. On optimal scheduling in duty-cycled industrial iot applications using ieee802.15.4e tsch. *IEEE Sensors Journal*, 13(10):3655–3666, Oct 2013.

[14] Maria Rita Palattella, Nicola Accettura, Luigi Alfredo Grieco, Gennaro Boggia, Mischa Dohler, and Thomas Engel. On optimal scheduling in duty-cycled industrial iot applications using ieee802. 15.4 e tsch. *IEEE Sensors Journal*, 13(10):3655–3666, 2013.

[15] M. Petrova, J. Riihijarvi, P. Mahonen, and S. Labella. Performance study of ieee 802.15.4 using measurements and simulations. In *Wireless communications and networking conference (WCNC)*, Wireless communications and networking conference (WCNC), pages 487–492, Las Vegas, USA, Apr 2006. IEEE.

[16] M. Qasem, A. Al-Dubai, I. Romdhani, B. Ghaleb, and W. Gharibi. A new efficient objective function for routing in internet of things paradigm. In *IEEE Conference on Standards for Communications and Networking (CSCN)*, IEEE Conference on Standards for Communications and Networking (CSCN), pages 1–6, Berlin, Germany, Oct 2016. IEEE.

[17] A. Raith and M. Ehrgott. A comparison of solution strategies for biobjective shortest path problems. *Computers & Operations Research*, 36:1299 ? 1331, 2009.

[18] S. Steiner and T. Radzik. Computing all efficient solutions of the biobjective minimum spanning tree problem. *Computers & Operations Research*, 35:198–211, 2008.

[19] J. Tripathi, J. De Oliveira, and J.P. Vasseur. Performance evaluation of the routing protocol for low-power and lossy networks (rpl). RFC 6687, 2012.

[20] Q. Wang, X. Vilajosana, and T. Watteyne. 6top protocol (6p). Internet-Draft, 2018.

[21] T. Watteyne, M. Palattella, and L. Grieco. Using ieee 802.15.4e time-slotted channel hopping (tsch) in the internet of things (iot): Problem statement. RFC 7554, 2015.

[22] A. Wierzbicki. Basic properties of scalarizing functionals for multiobjective optimization. *Mathematische Operationsforschung und Statistik, s. Optimization*, 8:55–60, 1977.

[23] M. Z. Win, P. C. Pinto, and L. A. Shepp. A mathematical theory of network interference and its applications. *Proceedings of the IEEE*, 97(2):205–230, Feb 2009.

[24] T. Winter, P. Thubert, A. Brandt, J. Hui, R. Kelsey, P. Levis, K. Pister, R. Struik, JP. Vasseur, and Alexander R. RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks. RFC 6550, 2012.

[25] L. D. Xu, W. He, and S. Li. Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4):2233–2243, Nov 2014.

[26] Dimitrios Zorbas, Vassilis Kotsiou, Fabrice Théoleyre, Georgios Z Papadopoulos, and Christos Douligeris. Lost: Localized blacklisting aware scheduling algorithm for ieee 802.15. 4-tsch networks. In *Wireless Days 2018*, Wireless Days 2018, pages 110–115, Dubai, UAE, 2018. IEEE.