# Energy-Optimal Data Aggregation and Dissemination for the Internet of Things

Fitzgerald, Emma; Pioro, Michal; Tomaszewski, Artur

*Published in:*
IEEE Internet of Things Journal

*DOI:*
10.1109/JIOT.2018.2803792

2018

*Document Version:*
Publisher's PDF, also known as Version of record

# Energy-Optimal Data Aggregation and Dissemination for the Internet of Things

Emma Fitzgerald* emma.fitzgerald@eit.lth.se
Michał Pióro*† michal.pioro@eit.lth.se
Artur Tomaszewski† artur@tele.pw.edu.pl

*Department of Electrical and Information Technology
Lund University
SE-221 00 Lund
Sweden

†Institute of Telecommunications
Warsaw University of Technology
Nowowiejska 15/19
00-665 Warsaw
Poland

***Abstract*—Established approaches to data aggregation in wireless sensor networks (WSNs) do not cover the variety of new use cases developing with the advent of the Internet of Things. In particular, the current push towards fog computing, in which control, computation, and storage are moved to nodes close to the network edge, induces a need to collect data at multiple sinks, rather than the single sink typically considered in WSN aggregation algorithms. Moreover, for machine-to-machine communication scenarios, actuators subscribing to sensor measurements may also be present, in which case data should be not only aggregated and processed in-network, but also disseminated to actuator nodes. In this paper, we present mixed-integer programming formulations and algorithms for the problem of energy-optimal routing and multiple-sink aggregation, as well as joint aggregation and dissemination, of sensor measurement data in IoT edge networks. We consider optimisation of the network for both minimal total energy usage, and min-max per-node energy usage. We also provide a formulation and algorithm for throughput-optimal scheduling of transmissions under the physical interference model in the pure aggregation case. We have conducted a numerical study to compare the energy required for the two use cases, as well as the time to solve them, in generated network scenarios with varying topologies and between 10 and 40 nodes. Although aggregation only accounts for less than 15% of total energy usage in all cases tested, it provides substantial energy savings. Our results show more than 13 times greater energy usage for 40-node networks using direct, shortest-path flows from sensors to actuators, compared with our aggregation and dissemination solutions.**

***Index Terms*—Internet of Things; fog computing; mixed-integer programming; sensor networks; aggregation**

## I. Introduction

Data aggregation has been a vital ingredient in wireless sensor network (WSN) design for some time, improving energy efficiency as nodes collate data on its way towards the sink, reducing the number of packets that need to be transmitted. However, we are now seeing the development of new architectures and use cases with the advent of the Internet of Things (IoT) and machine-to-machine (M2M) communication. Monitoring applications are still prevalent in application scenarios such as smart grids, vehicular communications, and smart cities, but are increasingly being complemented with actuation and in-network processing. The presence of actuators in the edge network means that data must not only be collected, but also disseminated to actuators that need it.

An example of such a use case could be a nuclear reactor with temperature sensors distributed throughout and actuators that can open valves to flood the reactor with water for emergency cooling in case of overheating. Each actuator needs to collect measurements from multiple temperature sensors, and if any of them exceeds a threshold, the valve is opened. In this case, the aggregation function would be to take the maximum of the aggregated temperature measurements. With the use of in-network processing, the functions to be executed at intermediate nodes can range from simple aggregation functions, such as the maximum function for the nuclear reactor, to more complex computations like compress-and-forward [1] or error correction, which may use an appreciable amount of energy in their execution.

Even in cases where data does not need to reach in-network actuators, but should simply be collected and transmitted to a single server, the expanding use of fog computing [2], [3] (also called network edge computing), and the increased heterogeneity of IoT networks means that the assumption of a single sink to collect this data is no longer valid. Rather, there may be multiple gateways with backbone connections located at the network edge and possibly operated by different parties. Much of the existing literature on data aggregation in WSNs has however focused on single-sink scenarios, and so there is a need to develop solutions that combine both gateway selection and routing for data aggregation for optimal energy efficiency. Fog computing not only brings new challenges, but also provides new opportunities. With more powerful nodes present at the network edge, greater coordination based on more complex, optimal algorithms becomes feasible in some cases, potentially reducing energy usage even further than existing approaches designed to be executed by resource-constrained sensor nodes. Even where the optimal solutions are not feasible to use in practice, for example in highly dynamic networks, it is nonetheless valuable to develop them in order to provide performance bounds for approximate solutions.

A further challenge that must be met in conjunction with the development of IoT systems is the rising density of networks, with many nodes in close proximity nd multiple networks often overlapping. This means that simple radio interference models, such as the unit disk or protocol models, may not be sufficient to capture the effects of second and higher order interference. This refers to interference stemming from multiple nodes transmitting simultaneously, even where any one of these nodes' transmissions would not be sufficient on its own to prevent successful decoding of packets at a given receiver node. Compatible sets (c-sets) [4] are capable of accounting for physical interference — that is, based on received signal strength — in a computationally efficient way, and have been successfully applied to modelling and optimisation of wireless mesh networks [5]–[9].

In this paper we present two new models for data aggregation in IoT networks. The first model, which we call $1K$, allows for collection of data at any of multiple, internet-connected gateways. Meanwhile, the second model, $nK$, is more applicable to scenarios with actuators present in the network, and provides for joint data aggregation and dissemination to all, or some minimum number of, multiple destination nodes. In both of these models, it is possible to specify a minimum number of sensor readings to be collected and aggregated. We then formulate these models as optimisation problems using mixed-integer programming (MIP), with the objective of minimum total energy usage, taking into account energy costs for both wireless transmissions and computation of aggregation functions. While network lifetime is also often a key performance metric, it requires a much more difficult optimisation problem to achieve. In this work, we also make an initial step towards network lifetime optimisation, by providing formulations for min-max per-node energy optimisation, for both the $1K$ and $nK$ scenarios.

Using our models, sensor and gateway selection, as well as the routing subgraphs for data aggregation and dissemination, are jointly optimised. For the $1K$ case, we also present a model and MIP formulation to optimise throughput of data collection, while maintaining minimum energy usage. We have performed numerical studies to investigate the minimum total energy and min-max energy required for data aggregation and dissemination, the maximum achievable throughput in the $1K$ case, and the time to solve network scenarios with varying numbers of nodes and topologies.

Our results show that aggregation constitutes a low energy cost, disproportionately so when taking into consideration the input energy costs for each transmission and aggregation operation. However, it provides a large advantage, with, in the most extreme case, 13.7 times higher energy usage using the most efficient paths between individual sensors and actuators, than using our $nK$ aggregation and dissemination solutions. Min-max energy optimisation provides overall less efficient solutions, with less aggregation, but with energy usage more fairly distributed amongst the nodes in the network.

The rest of this paper is organised as follows. In Section II we provide an overview of the related work in this area. Section III describes the network scenarios we consider, and Sections IV–VII detail our system model and formulations. In Section VIII we present the results of our numerical study. Finally, Section IX concludes this paper and discusses directions for future work.

## II. Related Work

The problem of energy-efficient data aggregation in wireless sensor networks first began to receive substantial attention in the 2000's [10]. Given the constrained computation capacity of sensor nodes and the distributed nature of such networks, much of the focus was on approximation algorithms for practical implementation, although some work has also been carried out on finding optimal solutions [1], [11]–[18]. Approximation algorithms for data aggregation can be mostly divided into tree-based [19]–[23], clustered [24]–[29], and hybrid [30] approaches [10]. Two of the most influential protocols stemming from such methods are LEACH [24], [25] and PEGASIS [20], on which much other work is based [31]–[35]. More recently, structure-free aggregation has been proposed [36], [37], which does not rely on the formation of a routing tree or clusters and is thus more suitable for highly mobile networks.

Most of this work assumes a single sink as the data collection point, and only data aggregation is performed; there is no dissemination to multiple actuators. Multiple sinks are considered in [19], however here there is still only one predetermined sink per data aggregation flow, making sink selection unnecessary. In [21], data dissemination (but not aggregation) is investigated, with multiple data sinks for each source, but each sink can only receive data from a single source. Optimal approaches with multiple sinks include [38], however here only cluster formation is optimised, not the overall routing tree, and cluster heads then transmit data to a central collection point using single-hop transmissions. In [17] multiple sinks were also used, giving an integer linear programming formulation for an optimal data collection schedule within each cluster such that at least one sensor measurement of each type is collected in each frame. Again, multihop transmission is not considered; sensor nodes transmit directly to their nearest collection point. Another integer linear programming solution, presented in [18], also has multiple gateways, however the problem considered here is that of gateway placement, rather than optimisation of a multihop aggregation tree.

A scenario with multiple gateways motivates the work in [1], but only a single sink is used in the actual optimisation formulation. This work is however interesting from another perspective, as it uses compress-and-forward to reduce the size of a node's transmission after it has overheard correlated measurements from other nodes. In this paper, we consider fixed packet sizes, and thus fixed transmission costs, however in M2M communications, it is common for measurements to be correlated, for example temporally, spatially, or semantically. The use of compress-and-forward would thus be a worthwhile direction to pursue in future iterations of our models.

In our $1K$ model, we address the optimisation of a multihop routing tree for data aggregation of at least some given minimum number of sensor measurements, and their propagation to any subset of multiple data collection nodes. As such, this constitutes joint gateway selection and routing. Moreover, if the total number of sensors in the network is larger than the number of required measurements, sensor selection is also included. In our $nK$ model, rather than routing data aggregates to any gateway, each destination node must receive measurements from a minimum number of sensors. To the best of the authors' knowledge, ours is the first work that considers such a joint aggregation and dissemination problem. Such models are necessary, since use cases for networks containing not only sensors producing data, but also actuators consuming this data, will become increasingly common in M2M communication scenarios [39].

We also provide throughput optimisation for the $1K$ case using a physical interference model based on compatible sets. Most existing work on data aggregation optimisation uses either a unit disk or protocol interference model, in which nodes have a larger interference range than their transmission ranges, but still only first order interferences from individual nodes are considered. Work that uses a more realistic physical interference model includes [11], [13], [14], [16], though with only a single sink and without joint aggregation and dissemination.

## III. System Overview

For our use case, we consider an Internet of Things network, with four types of nodes: sensor nodes, which produce data by taking measurements and can also transit and aggregate other sensors' measurements; aggregator nodes, which act as multihop relays and aggregators but do not take measurements; actuator nodes, which require sensor data in order to perform tasks; and gateways, which are fog nodes with backhaul connections. Fog nodes are thus capable of reaching a central server, such as a cloud service, via the Internet, and do not have the same energy constraints as other nodes. Sensors, actuators, and aggregators are battery-powered, and suffer energy costs for both wireless transmission and the computation required for data aggregation.

We thus have a wireless multihop network in which there is a set of data streams $\mathcal{S}$, representing different types of sensor measurement. In each stream, there are a number of origin nodes (sensors) $\mathcal{O}(s)$ producing data (one measurement from each sensor per frame), and a number of destination nodes (gateways or actuators) that wish to receive this data. Data may be aggregated by nodes during transmission such that a node receiving multiple packets belonging to a given stream can combine them according to some function (e.g., average, sum, count) and the node then only transmits a single packet representing this aggregate. Data from different streams is not aggregated; a stream is thus defined in a data-centric way as consisting of information that is subject to aggregation.

How the destination nodes receive the data depends on the application. In the simpler case, which we name $1K$, we wish to collect a certain number $K$ of (aggregated) measurements at some central server. In this case, the destination nodes represent fog gateways, and it is not important which gateway

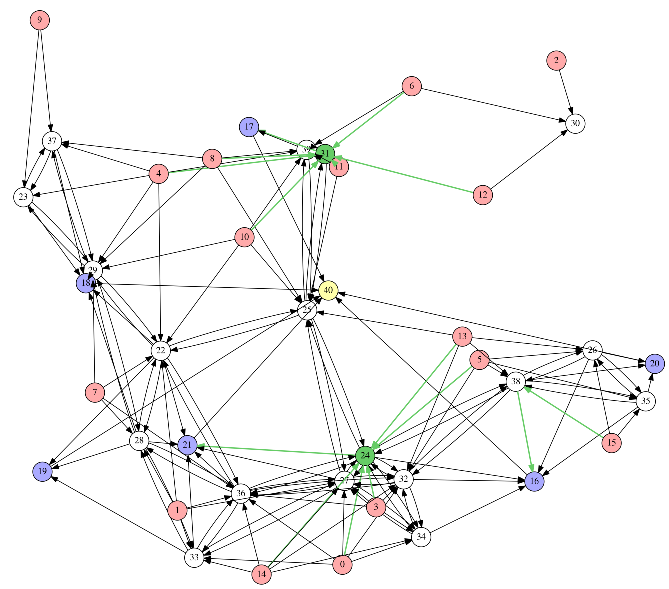

Fig. 1: An example $1K$ scenario, with $K = 12$. Sensors nodes are shown in red, aggregator nodes in white, and gateway nodes in blue. The central server is shown in yellow. Transit nodes and links belonging to the minimum-energy aggregation tree are highlighted in green.

collects any given measurement (Figure 1). In the figure, the transmission links between the gateways and the central server go over the Internet and thus do not incur energy costs. The second use case, $nK$, is where the destination nodes themselves wish to use the data, that is, the destination nodes are actuators that must perform some action based on the sensor measurements, and thus need to each collect a sufficient number of measurements (Figure 2). These measurements must originally come from different origin nodes, but multiple destination nodes may receive the same set of measurements.

The performance metrics of interest will also in general depend on the application, however here we will consider two objective functions: minimum total energy usage, and min-max energy usage. Energy is used both to transmit each packet, and to aggregate packets. This means there is an inherent trade-off in performing aggregation; each packet that must be aggregated costs energy, but this aggregation saves energy by reducing the number of packets to be transmitted.

Our first objective function aims to simply reduce the energy used by the network as a whole, while our second function targets network lifetime. Whether total energy usage or network lifetime is of greater interest depends on the specific application scenario. Where nodes are easily accessible for battery changes, or are connected to mains power (as in, for example, smart grids), total energy may be more important, whereas when nodes rely on a single battery or on energy harvesting, network lifetime may take precedence. To determine total energy usage requires a relatively simple additive function of the energy used by all nodes in the network. However,

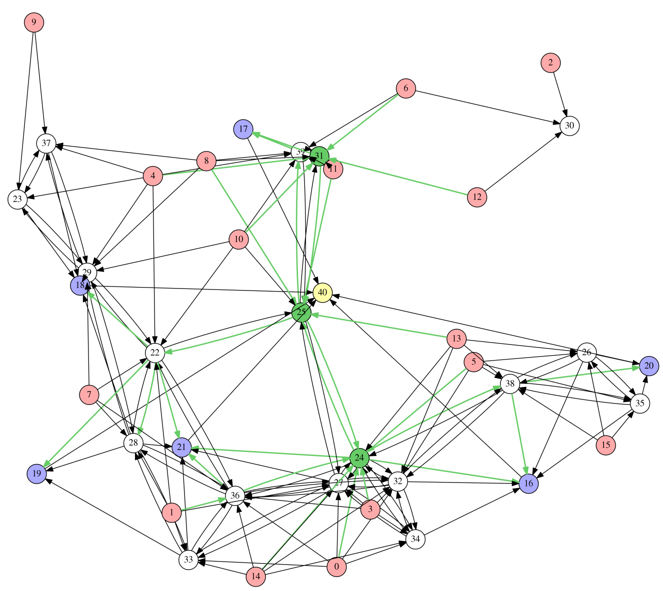

Fig. 2: An example $nK$ scenario., with $K = 12$, for the same network as in Figure 1. Sensor nodes are shown in red, aggregator nodes in white, and actuator nodes in blue. The central server, shown in yellow, is not used for the $nK$ case. Transit nodes and links belonging to the minimum-energy aggregation and dissemination subgraph are highlighted in green.

using network lifetime as the objective leads to much more complicated optimisation. This is because in this case we must consider which nodes are depleted of energy first, and how this affects the routing of the data streams. As such, the min-max energy optimisation we contribute in this work does not fully characterise network lifetime, but nonetheless provides the first step towards this goal.

The performance of data aggregation will also depend on the aggregation policy. For example, data may be aggregated based on temporal or geographical locality, or on data similarity. In our modelling, we take an agnostic approach to the aggregation policy. Different aggregation policies may be applied by selecting appropriate eligible origin, destination, and aggregator nodes, and this topology can then be supplied to our formulations. Our model allows for different sets of origin, destination and aggregator nodes for each data stream, giving fine-grained control over the data aggregation policy.

In the following sections, we present mixed-integer programming formulations for minimum total energy and min-max energy routing for both the $1K$ and $nK$ cases, as well as a formulation for minimum frame (maximum throughput) transmission for the $1K$ case.

## IV. $1K$ TOTAL ENERGY MINIMISATION

The method we employ for $1K$ aggregation is to use reverse arborescences (called r-arborescences in the following), that is, directed trees in which the arcs point towards the root node

| | |
|---|---|
| $\mathcal{V}$ | set of nodes (vertices) in the network |
| $\mathcal{E}$ | set of arcs $(v, w)$, $v, w \in \mathcal{V}$ indicating node $w$ is within transmission range of node $v$ (barring any interference) |
| $\mathcal{S}$ | set of data aggregation streams |
| $\mathcal{V}(s)$ | set of nodes belonging to the subgraph for stream $s \in \mathcal{S}$ |
| $\mathcal{E}(s)$ | set of arcs belonging to the subgraph for stream $s \in \mathcal{S}$ |
| $K(s)$ | number of packets that need to be collected to satisfy stream $s \in \mathcal{S}$ |
| $\mathcal{O}(s)$ | set of origin nodes for stream $s \in \mathcal{S}$ |
| $\mathcal{N}(s)$ | set of aggregator nodes for stream $s \in \mathcal{S}$ |
| $\mathcal{D}(s)$ | set of destination nodes for stream $s \in \mathcal{S}$ |
| $W$ | central server (sink node) |
| $\delta^-(v)$ | set of incoming arcs to node $v$ |
| $\delta^+(v)$ | set of outgoing arcs from node $v$ |
| $P$ | total transmission energy cost incurred by all nodes |
| $Q$ | total processing energy cost incurred by all nodes |
| $x_s^o$ | whether or not the packet in stream $s \in \mathcal{S}$ from node $o \in \mathcal{O}(s)$ is collected by the sink node $W$ |
| $y_{se}$ | whether or not arc $e \in \mathcal{E}$ is used for stream $s \in \mathcal{S}$ |
| $z_{se}^o$ | flow in stream $s \in \mathcal{S}$ from node $o \in \mathcal{O}(s)$ on arc $e \in \mathcal{E}$ |
| $g_{sv}$ | number of (aggregated) measurements in stream $s \in \mathcal{S}$ aggregated at node $v$ minus 1 (and 0 if there is no aggregation at $v$) |
| $B$ | transmission energy cost |
| $C$ | processing energy cost for aggregation |
| $\mathcal{A}(s)$ | optimal reverse arborescence for stream $s \in \mathcal{S}$ |
| $\mathcal{C}$ | set of all compatible sets (c-sets) for the network $(\mathcal{V}, \mathcal{E})$ |
| $T_c$ | number of time slots to be used by c-set $c \in \mathcal{C}$ |
| $n(e)$ | number of stream arborescences that use arc $e \in \mathcal{A}$ |
| $\pi_e$ | dual variables for frame minimisation |
| $\pi_e^*$ | optimal values of dual variables for frame minimisation |
| $Y_e$ | whether or not arc $e$ is present in the c-set generated by the pricing problem |
| $X_v$ | whether or not node $v$ is present in the c-set generated by the pricing problem |
| $\mathbb{B}$ | set of binary numbers$\{0, 1\}$ |
| $\mathbb{R}$ | set of real numbers |
| $\mathbb{R}_+$ | set of non-negative real numbers |

TABLE I: Summary of notation for $1K$ formulations.

such that all leaf nodes have exactly one path to the root node. Using r-arborescences ensures that no packet from an origin node in a given stream is counted twice, since it can only reach the central server along a single path. For a complete r-arborescence, each destination node for a stream should have an arc connecting it to the central server (the root node), however transmissions along these arcs are omitted from our energy model, and so they do not induce energy costs and are only used to ensure correct flows from origin to destination nodes. The notation used in the following formulations is summarised in Table I.

### A. Assumptions

For the $1K$ use case, a total of $K(s)$ measurements must be collected at destination nodes for each stream $s \in \mathcal{S}$. We assume that:

1) The network graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is composed of the set of nodes $\mathcal{V}$ and the set of (directed) arcs $\mathcal{E} \subseteq \mathcal{V}^2 \setminus \{(v, v) : v \in \mathcal{V}\}$. The set of arcs incoming to and the set of arcs outgoing from a given node $v \in \mathcal{V}$ (the so-called incoming and outgoing stars) are denoted by $\delta^-(v)$ and $\delta^+(v)$, respectively.
2) The set of nodes $\mathcal{V}$ is composed of three mutually disjoint subsets: the set of sensor nodes $\mathcal{O}$, the set of aggregator nodes $\mathcal{N}$, and the set of destination nodes

$\mathcal{D}$, and, additionally, of a central server (an artificial sink node) $W$. Thus, $\mathcal{V} = \mathcal{O} \cup \mathcal{N} \cup \mathcal{D} \cup \{W\}$.

3) Each destination node has only one outgoing arc, linking it to the central server: $\delta^+(d) = \{(d, W)\}, d \in \mathcal{D}$.
4) The only arcs incident to the central server are those from the destination nodes: $\delta^+(W) = \emptyset$, $\delta^-(W) = \{(d, W) : d \in \mathcal{D}\}$.
5) Origin nodes $o \in \mathcal{O}$ generate, transit, and aggregate packets. Aggregator nodes $v \in \mathcal{N}$ transit and aggregate packets, but do not generate their own. Destination nodes $d \in \mathcal{D}$ aggregate packets and forward them to the sink $W$. The sink then receives and aggregates these packets, but does not transmit any packets.
6) Each stream $s \in \mathcal{S}$ has a set of sensor nodes $\mathcal{O}(s)$ ($\mathcal{O}(s) \subseteq \mathcal{O}$), a set of aggregator nodes $\mathcal{N}(s)$ ($\mathcal{N}(s) \subseteq (N \cup \mathcal{O}) \setminus \mathcal{O}(s)$), and a set of destination nodes $\mathcal{D}(s)$ ($\mathcal{D}(s) \subseteq \mathcal{D}$). $\mathcal{O}(s)$, $\mathcal{N}(s)$, and $\mathcal{D}(s)$ are mutually disjoint. Note that all nodes in $N(s)$ act solely as aggregators for stream $s$, even though some of them (those in $O$) are origin nodes for other streams.
7) Packets of each stream $s \in \mathcal{S}$ have access only to the nodes in $\mathcal{V}(s) = \mathcal{O}(s) \cup \mathcal{N}(s) \cup \mathcal{D}(s) \cup \{W\}$, i.e., to the subgraph $\mathcal{G}(s) = (\mathcal{V}(s), \mathcal{E}(s))$ (where $\mathcal{E}(s) = \mathcal{E} \cap \mathcal{V}(s)^2$) of $\mathcal{G}$ generated by the set of nodes $\mathcal{V}(s)$.
8) For a given $s \in \mathcal{S}$, $\delta^-(s, v) = \delta^-(v) \cap \mathcal{E}(s)$ and $\delta^+(s, v) = \delta^+(v) \cap \mathcal{E}(s)$ are, respectively, the stars of the arcs incoming to and outgoing from node $v \in \mathcal{V}(s)$ in subgraph $\mathcal{G}(s)$.

### *B. Routing and aggregation optimisation*

The optimisation problem for minimum total energy routing and aggregation in the $1K$ case is formulated as follows:

$$\min \sum_{s\in\mathcal{S}} \big(P_s + Q_s\big) \tag{1a}$$

$$\sum_{o\in\mathcal{O}(s} x_s^o \geq K(s), \qquad s \in \mathcal{S} \tag{1b}$$

$$\sum_{e\in\delta^+(s,v)} z_{se}^o = \sum_{e\in\delta^-(s,v)} z_{se}^o, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s),\ v \in \mathcal{V}(s) \setminus \{o, W\} \tag{1c}$$

$$\sum_{e\in\delta^-(s,W)} z_{se}^o = x_s^o, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s) \tag{1d}$$

$$z_{se}^o \leq y_{se}, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s),\ e \in \mathcal{E}(s) \tag{1e}$$

$$\sum_{e\in\delta^+(s,v)} y_{se} \leq 1, \qquad s \in \mathcal{S},\ v \in \mathcal{V}(s) \tag{1f}$$

$$g_{sv} \geq \sum_{e\in\delta^-(s,v)} y_{se} - 1, \qquad s \in \mathcal{S},\ v \in \mathcal{N}(s) \tag{1g}$$

$$g_{so} \geq \sum_{e\in\delta^-(s,v)} y_{se} + x_s^o - 1, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s) \tag{1h}$$

$$P_s = B \sum_{e\in\mathcal{E}(s)\setminus\delta^-(s,W)} y_{se}, \qquad s \in \mathcal{S} \tag{1i}$$

$$Q_s = C \sum_{v\in\mathcal{O}(s)\cup\mathcal{N}(s)} g_{sv}, \qquad s \in \mathcal{S} \tag{1j}$$

$$x_s^o \in \mathbb{B}, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s) \tag{1k}$$

$$z_{se}^o \in \mathbb{R}_+, \qquad s \in \mathcal{S},\ o \in \mathcal{O}(s),\ e \in \mathcal{E}(s) \tag{1l}$$

$$y_{se} \in \mathbb{B}, \qquad s \in \mathcal{S},\ e \in \mathcal{E}(s) \tag{1m}$$

$$g_{sv} \in \mathbb{R}_+, \qquad s \in \mathcal{S},\ v \in \mathcal{O}(s) \cup \mathcal{N}(s) \tag{1n}$$

$$P_s, Q_s \in \mathbb{R}, \qquad s \in \mathcal{S}. \tag{1o}$$

The objective function (1a) seeks to minimise the total energy used by all nodes for both transmission and aggregation. For explaining the constraints, consider an arbitrary fixed data stream $s \in \mathcal{S}$. Then constraints (1b) ensure that sufficient sensor measurements are collected at the server to satisfy the stream. The flow conservation constraints (1c)–(1d) force that for each active sensor $o$ ($x_s^o = 1$), a flow of value 1 from $o$ to $W$ is realized along the arcs of $\mathcal{E}(s)$. Constraints (1e) make sure that any arc used by a flow is present in the r-arborescence $\mathcal{A}(s) = \{e \in \mathcal{E}(s) : y_{se} = 1\}$.

Constraints (1f) enforce one of the conditions characterizing the r-arborescence — at most one outgoing link to be used by each node in $\mathcal{V}(s)$. The remaining two r-arborescence conditions require that all active sensors are connected to $W$ within $\mathcal{A}(s)$, and that there are no loops in $\mathcal{A}(s)$. The first condition is enforced by the flow constraints (1c)–(1d) together with (1e), while the second condition is enforced by the objective that ensures that the number of arcs in an optimal $\mathcal{A}(s)$ is minimal.

Constraints (1g)–(1j) concern the energy costs. Constraints (1g) and (1h) induce aggregation processing costs only when a node receives (and generates, in the case of origin nodes) at least two packets to be aggregated. Note that the non-negativity of $g_{sv}$ is enforced by the range constraint (1n). Finally, constraint (1i) sums the transmission costs along all arcs of the r-arborescence $\mathcal{A}(s)$, and constraint (1j) sums the processing costs across all nodes in $\mathcal{A}(s)$.

### *C. Decomposition of the routing problem*

Note that the routing problem formulated in (1) can be decomposed to $|\mathcal{S}|$ separate minimisation subproblems solved for each stream $s \in \mathcal{S}$. Each such subproblem computes, for a given fixed $s$, the value $R(s) = \min\big(P_s + Q_s\big)$ with respect to constraints (1b)–(1o). Clearly, the minimum of (1a) will be equal to $R = \sum_{s\in\mathcal{S}} R(s)$ and the optimal solutions of the subproblems, when combined, will produce the optimal solution of (1).

The so-described decomposition is advantageous for time efficiency of the branch-and-bound solution algorithm since the solution time for the decomposed algorithm is linear with $|\mathcal{S}|$ while the solution time for (1) may scale more quickly with $|\mathcal{S}|$, due to the much larger number of variables and constraints in comparison to the subproblems..

### *D. Non-aggregating sensor nodes*

In formulation (1), we have assumed that all nodes are capable of aggregation and transiting (the latter function is not required for the server $W$). However, in some use cases, sensor nodes may be more limited and restricted to only broadcasting their own sensor measurements, without transiting or aggregating measurements generated by other sensors. Our formulation can also be applied to such a case. This is achieved simply by removing all the arcs incoming to the sensor nodes from the network graph. Note that then for all $s \in \mathcal{S}$ and $o \in \mathcal{O}(s)$, $\sum_{e\in\delta^-(s,o)} z^o_{se} = 0$ (because $\delta^-(s,o) = \emptyset$) so that the flow conservation constraints (1c)–(1d) remain valid. Note also that since the sensor nodes do not aggregate, variables $g_{so}$, $s \in \mathcal{S}$, $o \in \mathcal{O}(s)$, together with constraints (1h), are removed from the formulation. Additionally, the sets $\mathcal{O}(s)$ must be deleted from the summation range in (1j).

## V. Frame minimisation

Once the optimal reverse arborescences $\mathcal{A}(s) := \{e \in \mathcal{E}'(s) : y^*_{se} = 1\}$ (where $\mathcal{E}'(s) := \mathcal{E}(s) \setminus \delta^-(W)$ and $y^*$ is optimal for (1)) for all $s \in \mathcal{S}$ are found, the transmission pattern is optimised to achieve maximum throughput for the collection of sensor measurement streams. We adopt a similar method as in [8], based on the compatible sets (c-sets). A compatible set is a set of arcs that can successfully transmit simultaneously.

Consider a time-division multiple access (TDMA) frame consisting of a number of time slots, where each slot is sufficiently long for a node to transmit an (aggregated) measurement. We assume a TDMA-based medium access control scheme for this optimisation, since any other scheme will potentially have collisions and/or idle time, resulting in lower throughput. We thus obtain the maximum theoretical throughput, although in practice TDMA may not be the most suitable choice due to the coordination required.

We then wish to minimise the number of slots in the frame (the frame length) such that all necessary transmissions occur during a single frame. The frame then repeats to allow a stream of data to reach from the origin to the destination nodes. Note that an individual measurement does not traverse from origin to destination during a single frame, but rather each transmission required for this end-to-end route is provided for at least once in the frame. The pipelining of multiple subsequent measurements in the streams then fills all slots in the frame with transmissions after an initial warm-up period.

An appropriate optimisation problem can be constructed by means of the following integer programming (IP) formulation (where $\mathcal{A} := \bigcup_{s\in\mathcal{S}} \mathcal{A}(s)$).

$$\min \quad \sum_{c\in\mathcal{C}} T_c \tag{2a}$$

$$[\pi_e \geq 0] \quad \sum_{c\in\mathcal{C}(e)} T_c \geq n(e), \qquad e \in \mathcal{A} \tag{2b}$$

$$T_c \in \mathbb{Z}_+, \qquad c \in \mathcal{C}, \tag{2c}$$

where $n(e) := \sum_{s\in\mathcal{S}} y^*_{se}$ is the number of stream r-arborescences that use arc $e \in \mathcal{A}$. Here, the objective function (2a) minimises the total number of time slots needed in each frame, and constraints (2b) ensure that each c-set is active in a sufficient number of slots to transmit all packets on each arc.

To solve the above problem, we need to apply c-set generation (see [8]), since the number of all possible c-sets grows exponentially with the size of the network graph. For this we first solve the linear relaxation of (2) (i.e., assuming $T_c \in \mathbb{R}_+, c \in \mathcal{C}$) using the column generation approach. Suppose that an initial list of c-sets $\mathcal{C}$ is given (for example such a list can consist of all singletons $\{e\}, e \in \mathcal{A}$) and form the problem dual to (2) (in dual variables $\pi_e$, $e \in \mathcal{A}$):

$$\max \quad \sum_{e\in\mathcal{A}} \pi_e n(e) \tag{3a}$$

$$\sum_{e\in c} \pi_e \leq 1, \qquad c \in \mathcal{C} \tag{3b}$$

$$\pi_e \in \mathbb{R}_+, \qquad e \in \mathcal{A}. \tag{3c}$$

Consider a dual optimal solution $\pi^* = (\pi^*_e, e \in \mathcal{A})$ and suppose there exists a c-set, $c'$, say, outside the current list $\mathcal{C}$ with $\sum_{e\in c'} \pi^*_e > 1$. When $c'$ is added to the list ($\mathcal{C} := \mathcal{C} \cup \{c'\}$) then the new dual has one more constraint (3b) corresponding to $c'$, and this particular constraint is broken by the current optimal solution $\pi^*$. This means that the new dual polytope (for the updated c-set list) is a proper subset of the previous dual polytope, as the current optimal dual solution is cut off by the new dual constraints. Therefore, in the updated dual the maximum of (3a) cannot be greater than the previous maximum, and in fact in most cases it will be decreased. Since the maximum of the dual is equal to the minimum of the primal problem (2), adding the c-set $c'$ will usually decrease the frame length (this issue will soon be described in more detail). This procedure is repeated until there does not exist any such new c-set, which means that the final dual (and primal) solutions thus obtained are optimal even if $\mathcal{C}$ were extended to the list of all possible c-sets.

For generating the c-set $c'$ described above, the following pricing problem is used:

$$\max \quad \sum_{e\in\mathcal{A}} \pi^*_e Y_e \tag{4a}$$

$$\sum_{e\in\delta^-(v)\cup\delta^+(v)} Y_e \leq 1, \qquad v \in \mathcal{V}(\mathcal{A}) \tag{4b}$$

$$\sum_{e\in\delta^+(v)} Y_e = X_v, \qquad v \in \mathcal{V}(\mathcal{A}) \tag{4c}$$

$$\frac{1}{\gamma} p(v,w) Y_{(v,w)} \geq \big(\eta + \sum_{u\in\mathcal{V}(\mathcal{A})\setminus\{v,w\}} p(u,w) X_u\big) Y_{(v,w)}, \qquad (v,w) \in \mathcal{A} \tag{4d}$$

$$Y_e \in \mathbb{B}, \qquad e \in \mathcal{A} \tag{4e}$$

$$X_v \in \mathbb{B}, \qquad v \in \mathcal{V}(\mathcal{A}), \tag{4f}$$

where $\mathcal{V}(\mathcal{A})$ denotes the set of nodes that appear in the r-arborescences. The objective function (4a) generates a new c-set that maximizes the left-hand side of constraint (3b) for the current dual solution $\pi^*$. Constraints (4b) require that a node either be inactive or only transmit or receive on a single incoming or outgoing link at a time, and constraints (4c) ensure that a node is included in the generated c-set if it transmits (i.e., any of its outgoing arcs are in the c-set). The SINR constraints for successful transmission are expressed in (4d), where $\gamma$ is the SINR threshold for correct decoding of the data, $p(v,w)$ is the received power from node $v$ at node $w$, and $\eta$ is the noise power. To get an IP formulation, the bi-linearities in (4d) can be handled in the standard way by introducing auxiliary non-negative continuous variables $Z_{ue} = X_u \cdot Y_e$

| | |
|---|---|
| $x^{od}$ | whether or not the measurement from origin $o \in \mathcal{O}$ is received by destination $d \in \mathcal{D}$ |
| $z_e^{od}$ | flow of the measurement from origin $o \in \mathcal{O}$ to destination $d \in \mathcal{D}$ on arc $e \in \mathcal{E}$ |
| $y_e^o$ | whether or not arc $e \in \mathcal{E}$ carries the measurement from origin $o \in \mathcal{O}$ |
| $Y_e$ | whether or not arc $e \in \mathcal{E}$ carries an (aggregated) measurement |
| $X_v^{oo'}$ | whether or not the measurements from origins $o$, $o' \in \mathcal{O}$ are aggregated at node $v \in \mathcal{N} \cup \mathcal{D}$ |
| $G_v$ | energy required to broadcast from node $v \in \mathcal{O} \cup \mathcal{N}$ |
| $g_v$ | number of (aggregated) measurements aggregated at node $v \in \mathcal{N} \cup \mathcal{D}$ minus 1 (and 0 if there is no aggregation at $v$) |
| $u^o$ | whether or not the measurement from origin $o \in \mathcal{O}$ is received by any destination |
| $c(e)$ | transmission energy required on arc $e$ |
| $\mathcal{O}^{[2]}$ | set of all 2-element subsets of $\mathcal{O}$ |

TABLE II: Summary of notation for $nK$ formulation.

and additional constraints $Z_{ue} \leq X_u$, $Z_{ue} \leq Y_e$, $Z_{ue} \geq X_u + Y_e - 1$. If the maximum of (4) is greater than 1 then the c-set $c' := \{e \in \mathcal{A} : Y_e^* = 1\}$ (where $Y^*$ is an optimal solution) is added to the current list of the c-sets used in (2).

The pricing problem will produce a c-set $c'$ (if any) whose constraint (3b) is maximally broken by $\pi^*$. If the primal problem (the linear relaxation of (2)) is solved by the simplex algorithm, then when a new column (variable $T_{c'}$) is added to (2) and this column replaces one of the current basic variables, the local rate of decrease in the value of the primal objective (2a) will be equal to $1 - \sum_{e \in c'} \pi_e^*$ (in the simplex method this value is called the reduced cost of variable $T_{c'}$), and thus maximal over all c-sets. After column $c'$ enters the basis, the value of (2a) will be decreased by $(1 - \sum_{e \in c'} \pi_e^*)t$, where $t$ is the value assigned to variable $T_{c'}$ by the simplex pivot operation. Certainly, if the current basic solution is degenerate then $t$ may have to stay at the zero value, and in effect the objective function will not be decreased. Nevertheless, adding variable $T_{c'}$ to the problem is necessary for the simplex algorithm to proceed towards the optimal vertex solution. Observe that the final simplex solution of the linear relaxation of (2) will contain at most $|\mathcal{A}|$ non-zero values in the vector $T^* = (T_c^*, c \in \mathcal{C})$ specifying the optimal vertex. As shown in [40], this property helps to develop a reasonable heuristic for solving the primal problem.

After generating the c-set list $\mathcal{C}$, problem (2) is solved by the branch-and-bound algorithm in integer variables $T_c$, $c \in \mathcal{C}$. Note that the integer solution, $T_c^*$, $c \in \mathcal{C}$, obtained thereby may in general be suboptimal, since there can exist c-sets that are not necessary to solve the linear relaxation but are required for achieving the true optimum of (2) (when all c-sets are considered). In fact, to assure such true optimality, the IP problem should be solved using the branch-and-price algorithm [41] instead of the price-and-branch algorithm (i.e., the algorithm described above). Note also that a reasonable suboptimal solution of (2) can be obtained by simply rounding up the (fractional) linear relaxation solution.

## VI. $nK$ TOTAL ENERGY MINIMISATION

For the $nK$ case, we consider only one data aggregation stream and therefore we omit index $s$. Additional notation introduced for the $nK$ formulation is summarised in Table II.

### A. Assumptions

In the $nK$ case, at least $K$ measurements must be collected at each destination node. We assume that:

1) The set of nodes $\mathcal{V}$ is composed of three mutually disjoint subsets, the set of sensor (origin) nodes $\mathcal{O}$, the set of aggregator nodes $\mathcal{N}$, and the set of destination nodes $\mathcal{D}$. Thus, $\mathcal{V} = \mathcal{O} \cup \mathcal{N} \cup \mathcal{D}$.
2) Origin nodes generate, transit, and aggregate packets. Aggregation nodes $\mathcal{N}$ transit and aggregate packets. Destination nodes $\mathcal{D}$ can aggregate packets but do not transit them (hence, $\delta^+(d) = \emptyset$, $d \in \mathcal{D}$).
3) In effect, all aggregates (the measurement broadcast from a sensor node is also treated as an aggregate) received at a node in $\mathcal{O} \cup \mathcal{N}$ are collected, aggregated and broadcast.

### B. Routing, aggregation, and dissemination optimisation

Below we present an IP formulation of routing optimisation for aggregation and dissemination of data in the $nK$ case.

$$\min \quad P + Q \tag{5a}$$
$$\sum_{o \in \mathcal{O}} x^{od} \geq K, \qquad d \in \mathcal{D} \tag{5b}$$
$$\sum_{e \in \delta^+(v)} z_e^{od} = \sum_{e \in \delta^-(v)} z_e^{od}, \qquad o \in \mathcal{O},\ d \in \mathcal{D},\ v \in \mathcal{V} \setminus \{o, d\} \tag{5c}$$
$$\sum_{e \in \delta^-(d)} z_e^{od} = x^{od}, \qquad o \in \mathcal{O},\ d \in \mathcal{D} \tag{5d}$$
$$z_e^{od} \leq Y_e, \qquad o \in \mathcal{O},\ d \in \mathcal{D},\ e \in \mathcal{E} \tag{5e}$$
$$Y_e \leq \sum_{o \in \mathcal{O}} \sum_{d \in \mathcal{D}} z_e^{od}, \qquad e \in \mathcal{E} \tag{5f}$$
$$z_e^{od} \leq y_e^o, \qquad o \in \mathcal{O},\ d \in \mathcal{D},\ e \in \mathcal{E} \tag{5g}$$
$$y_e^o \leq \sum_{d \in \mathcal{D}} z_e^{od}, \qquad o \in \mathcal{O},\ e \in \mathcal{E} \tag{5h}$$
$$\sum_{e \in \delta^-(v)} y_e^o \leq 1, \qquad o \in \mathcal{O},\ v \in \mathcal{V} \tag{5i}$$
$$X_v^{oo'} \geq y_e^o + \Big( \sum_{e' \in \delta^-(v) \setminus \{e\}} y_{e'}^{o'} \Big) - 1, \qquad v \in \mathcal{V}, e \in \delta^-(v),\ \{o, o'\} \in \mathcal{O}^{|2|} \tag{5j}$$
$$\sum_{v \in \mathcal{V}} X_v^{oo'} \leq 1, \qquad \{o, o'\} \in \mathcal{O}^{|2|} \tag{5k}$$
$$g_v \geq \sum_{e \in \delta^-(v)} Y_e - 1, \qquad v \in \mathcal{N} \cup \mathcal{D} \tag{5l}$$
$$g_o \geq \sum_{e \in \delta^-(v)} Y_e + u^o - 1, \qquad o \in \mathcal{O} \tag{5m}$$
$$u^o \geq x^{od}, \qquad o \in \mathcal{O},\ d \in \mathcal{D} \tag{5n}$$
$$G_v \geq b(e) Y_e, \qquad v \in \mathcal{O} \cup \mathcal{N},\ e \in \delta^+(v) \tag{5o}$$
$$P = \sum_{v \in \mathcal{O} \cup \mathcal{N}} G_v, \tag{5p}$$
$$Q = C \sum_{v \in \mathcal{V}} g_v \tag{5q}$$
$$x^{od} \in \mathbb{B}, \qquad o \in \mathcal{O},\ d \in \mathcal{D} \tag{5r}$$

$$z_e^{od} \in \mathbb{B}, \qquad o \in \mathcal{O},\, d \in \mathcal{D},\, e \in \mathcal{E} \quad (5s)$$
$$u^o \in \mathbb{B}, \qquad o \in \mathcal{O} \quad (5t)$$
$$y_e^o \in \mathbb{R}_+, \qquad o \in \mathcal{O},\, e \in \mathcal{E} \quad (5u)$$
$$Y_e \in \mathbb{R}_+, \qquad e \in \mathcal{E} \quad (5v)$$
$$X_v^{oo'} \in \mathbb{R}_+, \qquad v \in \mathcal{N} \cup \mathcal{D},\, \{o, o'\} \in \mathcal{O}^{|2|} \quad (5w)$$
$$G_v \in \mathbb{R}_+, \qquad v \in \mathcal{O} \cup \mathcal{N} \quad (5x)$$
$$g_v \in \mathbb{R}_+, \qquad v \in \mathcal{V}. \quad (5y)$$

As in the $1K$ case, the objective function (5a) minimises the total energy usage for transmission and aggregation. Constraints (5b) ensure each destination node receives a sufficient number of sensor measurements. The flow conservation constraints (5c)–(5d) require there to be a flow from each origin node $o$ to each destination node $d$ that is supposed to collect the origin node's measurement ($x^{od} = 1$). Constraints (5e)–(5h) require arcs to be used if and only if they carry flows. Note that in the optimal solution variables $Y_e$ will be binary as they are forced to 1 by binary variables $z_e^{od}$ on arcs used for transmission (constraints (5e)), while on arcs with no transmissions they are forced to 1 (constraints (5f)). Variables $y_e^o$ have the same property, since they are either forced to 1 by constraints (5g), or to zero by constraints (5h).

Constraints (5i)–(5k) provide the conditions for aggregation. In particular, constraints (5i)) ensure that no measurement may be received on more than one arc incoming to a node. Next, constraints (5j) force variable $X_v^{oo'}$ to be equal to 1 when measurements from two different origins $o$ and $o'$ enter node $v$ on two different arcs and thus are aggregated at $v$. Finally, constraints (5k) make sure that two packets from different origin nodes can be aggregated at most once.

Any node that receives at least two packets aggregates them, and will incur a processing cost proportional to the number of packets aggregated, given by constraints (5l)–(5n). Constraint (5o) sets each node's transmission cost to the highest cost of any of the outgoing links on which it broadcasts a packet. Lastly, constraints (5p) and (5q) calculate the total transmission and processing costs, respectively.

**Remark 1:** Note that the energy calculation in (5p) is optimistic, as, due to interference, it may take more than one broadcast to deliver an aggregate from a node $v$ to all neighbouring nodes in $\{w : (v, w) \in \delta^+(v), Y_{(v,w)} = 1\}$. This issue must be solved during transmission scheduling. Note that this problem is not present in the $1K$ case, since there all transmissions are unicast.

We do not, however, consider frame minimisation for the $nK$ case here. Due to the multicast transmissions for dissemination of measurements to multiple destinations, frame minimisation becomes significantly more complicated than for the $1K$ case (see [42]). As such, we leave this problem for future work.

**Remark 2:** The formulation provided above can be adapted to multiple streams in a similar manner as in Section IV. The formulation as given here then becomes a subproblem solved on the subgraph for the corresponding stream, and the overall objective becomes the sum of the objectives of the subproblems for each stream.

**Remark 3:** The case where sensor nodes do not aggregate or transit data can be accommodated for the $nK$ case in a similar way to the $1K$ case. Constraints (5m) and (5n), along with variables $u^o, o \in \mathcal{O}$, should be removed, and the summation range in constraints (5q) changed from $\mathcal{V}$ to $\mathcal{N} \cup \mathcal{D}$.

## VII. Min-max energy optimisation

In order to work towards optimisation for the network lifetime, as an alternative to total energy usage, we here consider min-max energy optimisation, that is, optimising for the minimum energy usage of the node that uses the most energy. This is the node that will deplete its battery and fail first, so this gives a measure of the time to failure of the network. For a complete characterisation of network lifetime, routes should then be re-computed after the failure of this node, using the min-max energy of the remaining nodes as the objective function. This process would then be iterated until the remaining set of nodes yields an infeasible problem instance, at which point it is no longer possible to find a routing solution, and the network can be considered to have ultimately failed.

For min-max energy optimisation for the $1K$ case, we present the following extension based on formulation (1):

$$\min\ Z \quad (6a)$$
$$Z \geq G_v, \qquad v \in \mathcal{O} \cup \mathcal{N} \quad (6b)$$
$$G_v = \sum_{s \in \mathcal{S}} G_{sv}, \qquad v \in \mathcal{N} \quad (6c)$$
$$G_{sv} = Cg_{sv} + p \sum_{e \in \delta^+(v)} y_{se} + q \sum_{e \in \delta^-(v)} y_{se}, \qquad v \in \mathcal{O}(s) \cup \mathcal{N}(s),\, s \in \mathcal{S} \quad (6d)$$
$$Z, G_{sv}, G_v \in \mathbb{R}_+, \qquad s \in \mathcal{S},\, v \in \mathcal{V}. \quad (6e)$$

Above, $p + q = B$ where $p$ is the cost of broadcasting a packet from a node and $q$ is the cost of receiving a packet at a node. The new objective and constraint (6b) substitute the old objective (1a), and constraints (6c)–(6d) substitute constraints (1j) and (1i). The minimum of $Z$ expresses the energy consumed by the node with the maximum consumption during one measurement cycle.

For the $nK$ case (formulation (5)) again the objective (6a) replaces the old objective (5a), and we additionally replace constraints (5p) and (5q) with the following.

$$Z \geq H_v, \qquad v \in \mathcal{O} \cup \mathcal{N} \quad (7a)$$
$$H_v = Cg_v + pG_v + q \sum_{e \in \delta^-(v)} Y_e, \qquad v \in \mathcal{O} \cup \mathcal{N} \quad (7b)$$
$$H_v \in \mathbb{R}_+, \qquad v \in \mathcal{O} \cup \mathcal{N} \quad (7c)$$

## VIII. Numerical study

We conducted a numerical study with network instances with 10 to 40 nodes, generated using [43]. Nodes were placed

| Nodes | Area width [m] | $K$ | $\lvert\mathcal{O}\rvert$ | $\lvert\mathcal{D}\rvert$ | $\mathcal{N}$ |
|---|---|---|---|---|---|
| 10 | 122.47 | 3 | 4 | 2 | 4 |
| 15 | 150.0 | 5 | 6 | 3 | 6 |
| 20 | 173.21 | 6 | 8 | 3 | 9 |
| 25 | 193.65 | 8 | 10 | 4 | 9 |
| 30 | 212.13 | 9 | 12 | 5 | 13 |
| 35 | 229.13 | 11 | 14 | 6 | 15 |
| 40 | 244.95 | 12 | 16 | 6 | 18 |

TABLE III: Parameters used for the numerical study.

uniformly randomly in a square area, and a subset of them were designated as origin and destination nodes. The area was scaled with the number of nodes to maintain a constant network density. The number of measurements to collect, number of sensors and number of destination nodes were also scaled with the total number of nodes in the network. However, due to quantisation, this scaling is not exact, but rather follows the relation $\lceil r|\mathcal{V}|\rceil$, where $r = 0.4$ for the number of sensors, $r = 0.15$ for the number of destination nodes, and $r = 0.75$ for the number of measurements to collect. These numbers were chosen in order to generate network scenarios where aggregation and routing were required to a sufficient extent to illustrate the performance of our models. Generated networks for which one or more of the optimisation problems was infeasible, for example due to disjoint network components, were discarded and a new instance was generated instead. The results described below are obtained from twenty network instances for each data point. A summary of the parameters used is shown in Table III. These parameters are used for all experiments, both those concerning $1K$ optimisation, and those for $nK$ optimisation. For both the $1K$ and $nK$ cases, one data aggregation stream was used.

Links between nodes were established based on a transmission power of 20 mW, signal attenuation (path loss) proportional to distance with an exponent of 4, noise power of -81 dBm, and a SINR threshold of 8 dB. These values were also used for the SINR calculations in the frame minimisation for the $1K$ case. The transmission energy cost and aggregation processing cost were set to 5 and 1, respectively. Note that the absolute values of these costs are unimportant for the optimisation; rather, it is the ratio of the two that matters.

### *A. Total energy cost*

Our main experiments concerned the total energy cost using formulations (1) and (5). The energy costs for $1K$ routing are shown in Figure 3, and for $nK$ routing in Figure 4. In the figures, the average energy usage across all twenty network instances is shown for each data point, along with 95% confidence intervals. The bulk of the energy used is for transmission. Although this is of course influenced by the choice of the relative costs, transmission energy nonetheless accounts for a disproportionately large amount of the overall energy used. For small networks (10 nodes), the energy used for the $1K$ and $nK$ cases is similar, but for $nK$ we see a faster increase as the network grows larger.

Figures 5 and 6 show the number of nodes that aggregate packets for the $1K$ and $nK$ cases, respectively. In both cases, the number of aggregating nodes increases steadily with the number of nodes in the network, although there is greater variance for the $nK$ case. This indicates that aggregation is being used in the routing solutions proportionally with the network size. Examining the solutions produced reveals that this is because aggregation generally occurs close to the origin nodes, which will lower transmission costs as then only a single aggregate packet needs to be transmitted further across the network. Example solutions are shown in Figures 7 and 8, illustrating this behaviour.

Boxplots of the solution times for routing in $1K$ and $nK$ scenarios are shown in Figures 9 and 10, respectively. Note that solution times are shown with a logarithmic scale due to the large range of solution times observed. For all the network scenarios generated, the solution times for the $1K$ case remain small (on the order of seconds or tens of seconds), although there is greater variance as the network size increases. The solution time for the $nK$ case increases much more rapidly, soon becoming impractical except for static network scenarios. Nonetheless, our results can be used to provide performance bounds for approximation algorithms. For the $nK$ scenarios, there is even greater variance in the solution times than for $1K$. Further work is needed to investigate which network

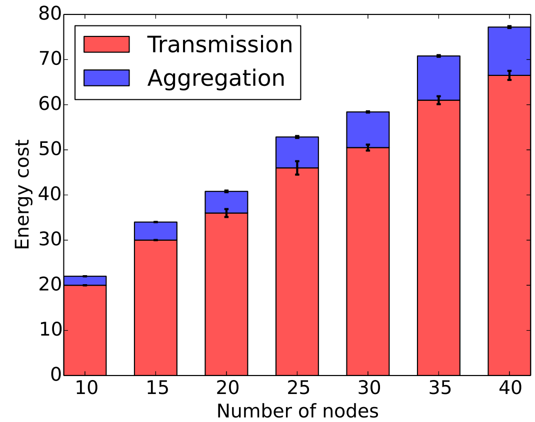


Fig. 3: Mean energy costs for the $1K$ case, with 95% confidence intervals.

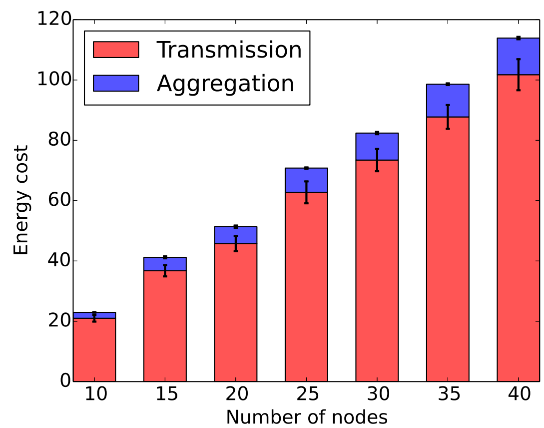


Fig. 4: Mean energy costs for the $nK$ case, with 95% confidence intervals.

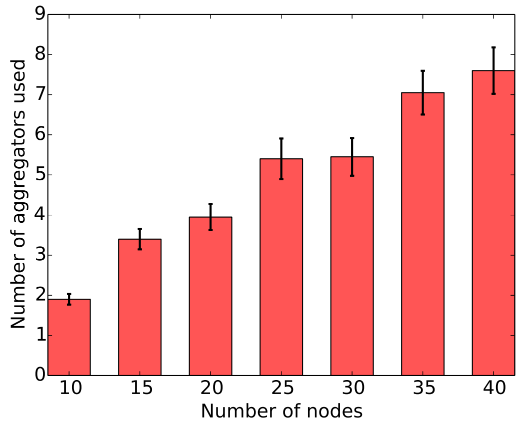


Fig. 5: Mean number of aggregating nodes used for the $1K$ case, with 95% confidence intervals.

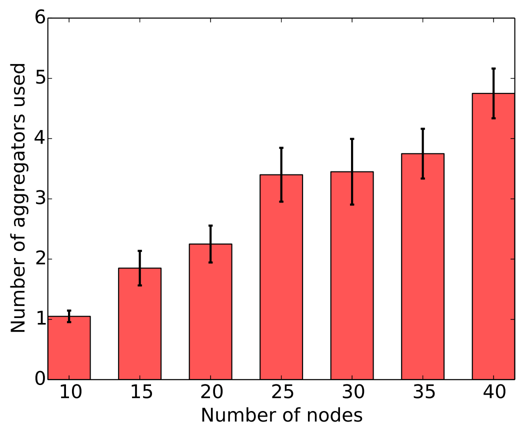


Fig. 6: Mean number of aggregating nodes used for the $nK$ case, with 95% confidence intervals.

characteristics contribute to a shorter or longer solution time. All problem instances were solved on an Intel Core i7-3770K CPU (3.5 GHz) with 8 virtual cores (4 cores with 2 threads each), and 8 GB RAM.

The results for the frame minimisation in $1K$ scenarios are shown in Table IV. Average solution times for frame minimisation were less than 0.5 s for all network sizes, with little variance, and the frame length grew only slowly with the number of nodes in the network. This indicates that frame minimisation does not add substantially to the work required to find a solution, and should thus be feasible in any scenario where aggregation routing is feasible.

### *B. Energy Costs Without Aggregation*

While aggregation energy costs represent a relatively small proportion of the overall energy costs, the ability to aggregate data is critical to energy efficiency. To illustrate this, we here present results showing the energy costs without aggregation. For these results, the scenarios were generated as in the previous section, however, routing was performed using a more traditional shortest path approach. Specifically, energy costs were computed using the following steps for $1K$ aggregation.

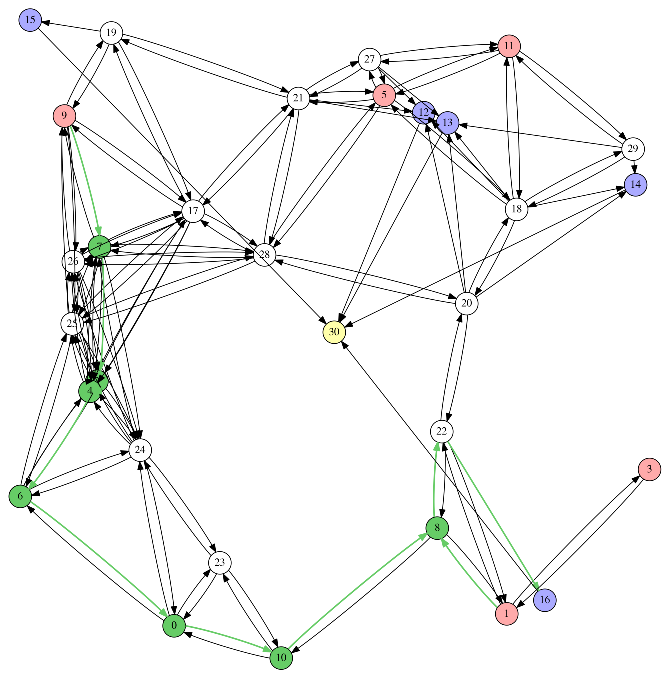

Fig. 7: An example routing solution for $1K$ with 30 nodes. Used links and aggregating nodes are shown in green.

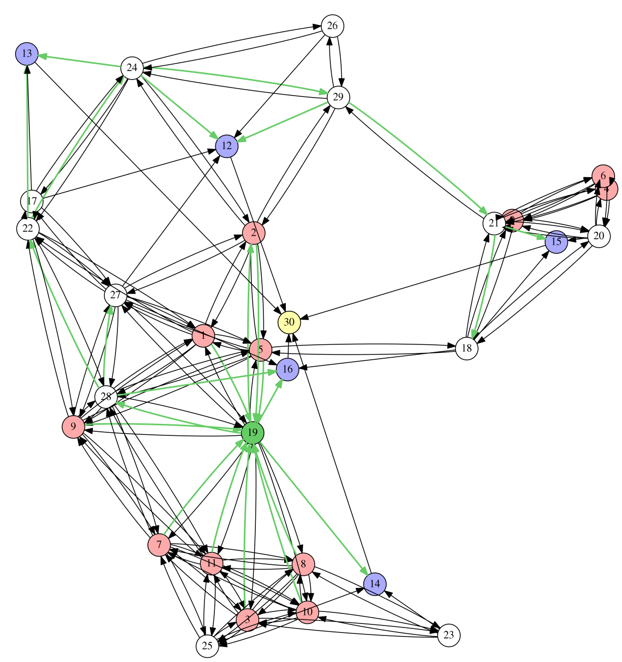

Fig. 8: An example routing solution for $nK$ with 30 nodes. Used links and aggregating nodes are shown in green.

1) For each stream $s \in \mathcal{S}$, and for each origin node $o \in \mathcal{O}(s)$, find a shortest path from $o$ to the central node $W$ in subgraph $\mathcal{G}(s)$.
2) For each stream $s \in \mathcal{S}$, use $K(s)$ shortest paths from origins in $\mathcal{O}(s)$, out of $|\mathcal{O}(s)|$ candidate paths.
3) For each arc $e \in \mathcal{E}$, compute $n(e)$, that is, how many of

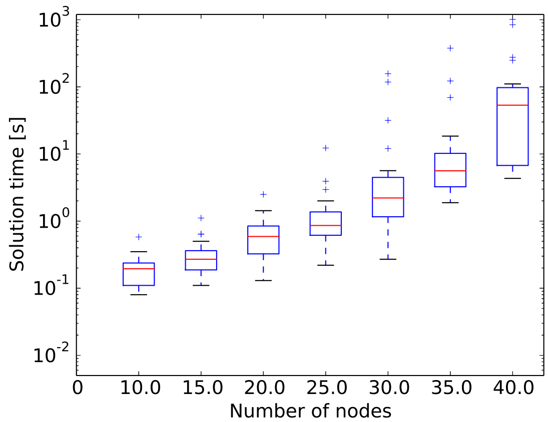


Fig. 9: Solution time for $1K$ routing, shown in logarithmic scale. The red lines indicate the median, and the blue boxes give the first and third quartiles.

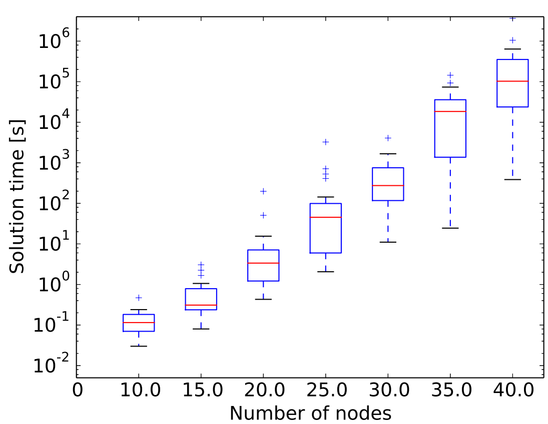


Fig. 10: Solution time for $nK$ routing, shown in logarithmic scale. The red lines indicate the median, and the blue boxes give the first and third quartiles.

the selected paths (for all streams) use arc $e$.

4) The (minimal) energy cost is then $\sum_{e \in \mathcal{E}} Bn(e)$.

Shortest paths were computed using the NetworkX library [44]. For the $nK$ case, the procedure was similar, except that shortest paths were computed between each pair of origin and destination nodes, and then the $K$ shortest paths were chosen to each destination.

| N | Frame Length | Solution Time [s] | Iterations | C-sets Generated |
|---|---|---|---|---|
| 10 | 3.0 (±0.14) | 0.04 (±0.01) | 2.9 (±0.24) | 5.9 (±0.24) |
| 15 | 3.2 (±0.22) | 0.09 (±0.03) | 5.55 (±0.64) | 10.55 (±0.64) |
| 20 | 3.4 (±0.29) | 0.09 (±0.02) | 6.5 (±0.69) | 12.7 (±0.68) |
| 25 | 3.7 (±0.24) | 0.13 (±0.03) | 7.75 (±0.68) | 15.95 (±0.78) |
| 30 | 4.3 (±0.42) | 0.14 (±0.03) | 8.9 (±0.84) | 18.0 (±0.89) |
| 35 | 4.4 (±0.35) | 0.27 (±0.05) | 10.9 (±0.99) | 22.1 (±1.05) |
| 40 | 4.6 (±0.35) | 0.27 (±0.05) | 10.1 (±0.95) | 22.4 (±1.01) |

TABLE IV: Frame minimisation results for $1K$ scenarios. Values shown are means across the 20 generated scenarios for each number of nodes, with 95% confidence intervals in parentheses. N = number of nodes.

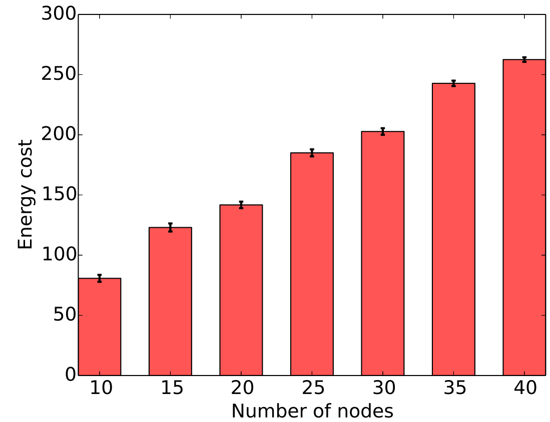


Fig. 11: Mean energy costs without aggregation for the $1K$ case, with 95% confidence intervals.

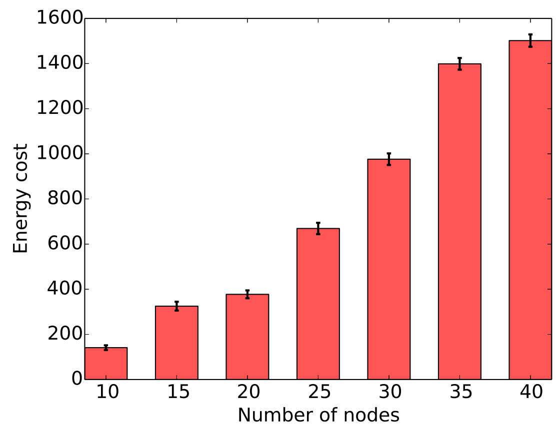


Fig. 12: Mean energy costs without aggregation for the $nK$ case, with 95% confidence intervals.

Figures 11 and 12 show the energy costs without aggregation for the $1K$ and $nK$ cases, respectively. The savings in energy usage are substantial in both cases, indeed, for the $nK$ case we see an order of magnitude lower energy cost. This illustrates the importance of finding efficient routing solutions for data aggregation, especially for emerging M2M scenarios in which dissemination of aggregated data to actuator nodes is also required.

### C. Min-Max Energy

In the following, we present results for min-max energy optimisation. In this case the objective function is the minimum energy usage of the node that uses the most energy, and as such, the results are not directly comparable to those for total energy cost optimisation. This is because here we obtain the energy usage of a single node (the node that uses the most energy), whereas for the total energy cost case we obtain the energy usage for the entire network. Since energy costs for aggregation and transmission are combined in the min-max formulations (6) and (7), we only show the total energy usage, rather than a breakdown of the energy used for

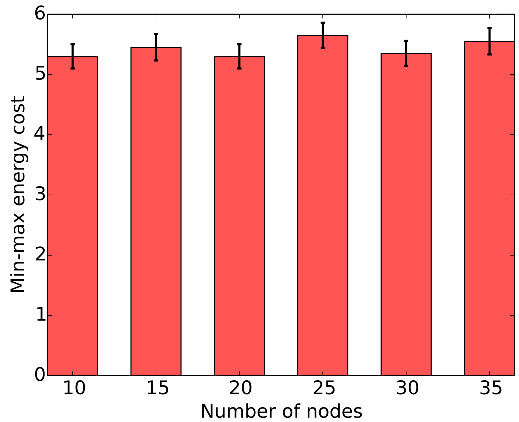


Fig. 13: Mean energy costs for the $1K$ case with min-max energy optimisation, with 95% confidence intervals.

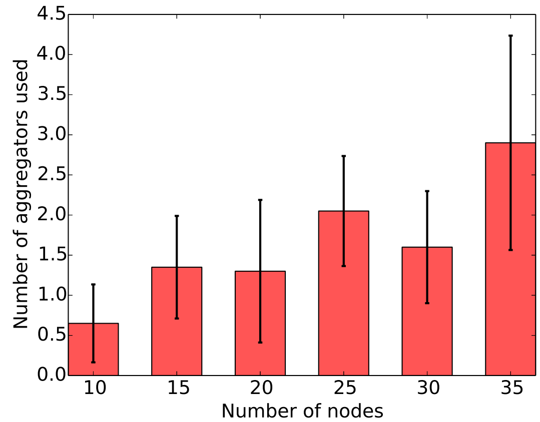


Fig. 15: Mean number of aggregating nodes used for the $1K$ case with min-max energy optimisation, with 95% confidence intervals.

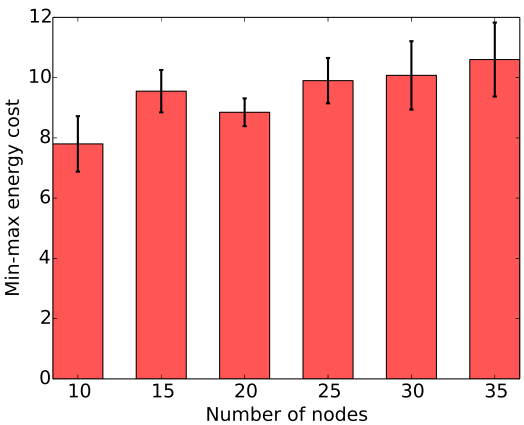


Fig. 14: Mean energy costs for the $nK$ case with min-max energy otimisation, with 95% confidence intervals.

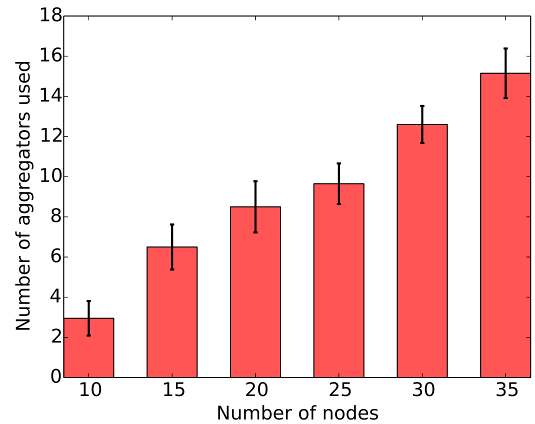


Fig. 16: Mean number of aggregating nodes used for the $nK$ case with min-max energy optimisation, with 95% confidence intervals.

these two functions as previously. We present here only results for networks up to a maximum of 35 nodes, rather than 40 nodes as previously, due to the excessive execution times for the latter case.

Figures 13 and 14 show the min-max energy cost for the $1K$ and $nK$ cases, respectively. Compared with the total energy cost, here we see little variation in the energy cost as the number of nodes increases. Since the total number of packets to be collected does increase with the number of nodes, this suggests that the work to aggregate and forward data packets is spread throughout the network, rather than disproportionately allocated to one or a few bottleneck nodes. This is a positive result in terms of the overall network lifetime, as in general network lifetime is improved by more fairly distributing energy-draining work amongst the nodes in the network. However, a full investigation of network lifetime is required to properly demonstrate the performance in this case.

The number of nodes that aggregate is shown in Figures 15 and 16. For the $1K$ case, the number of aggregating nodes is much lower than for total energy optimisation, and does not display the same relationship with the number of nodes in the network. This is because although aggregation in general saves energy (by reducing transmission energy costs), it nonetheless represents an energy cost for the individual aggregating node. For min-max optimisation therefore, where individual node energy is more important than overall energy, aggregation is actually avoided unless it reduces redundant transmission for the specific aggregating node itself. This results in relatively few or even no nodes aggregating, and overall more paths used throughout the network such that each node also only transmits one or a few times. An example solution is shown in Figure 17. In this solution, no aggregation occurs at all, and the origin nodes' measurements take disjoint paths to the destination nodes.

For the $nK$ case, on the other hand, since more packets need to be transmitted overall, it is harder for measurements to take disjoint paths. This means that aggregation becomes necessary in order to save energy at individual nodes. Here, we see a larger number of aggregating nodes than for total energy optimisation. This is because, although aggregation occurs in both cases, for min-max energy usage it is more distributed, with each node typically only aggregating few packets. This

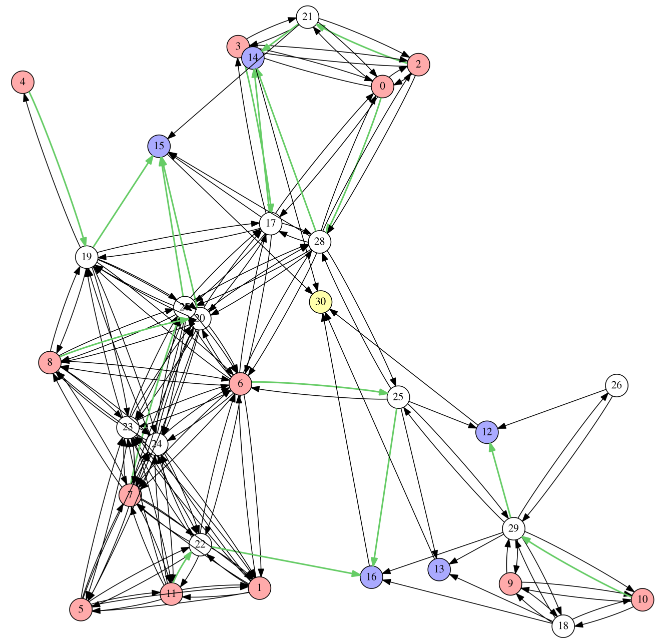

Fig. 17: An example routing solution for $1K$ with min-max energy optimisation, with 30 nodes. Used links and aggregating nodes are shown in green.

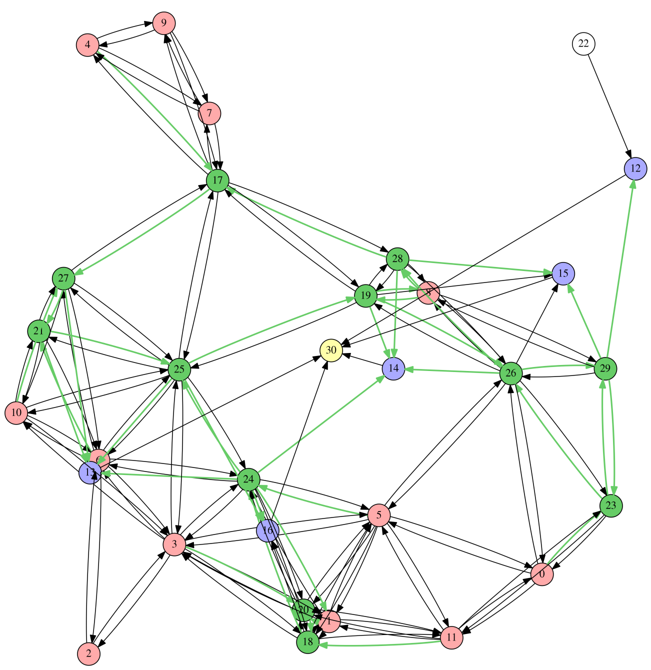

Fig. 18: An example routing solution for $nK$ with min-max energy optimisation, with 30 nodes. Used links and aggregating nodes are shown in green.

can be seen in the example solution in Figure 18, where both aggregating nodes and links used are distributed widely throughout the network.

The routing solution times are shown in Figure 19 for

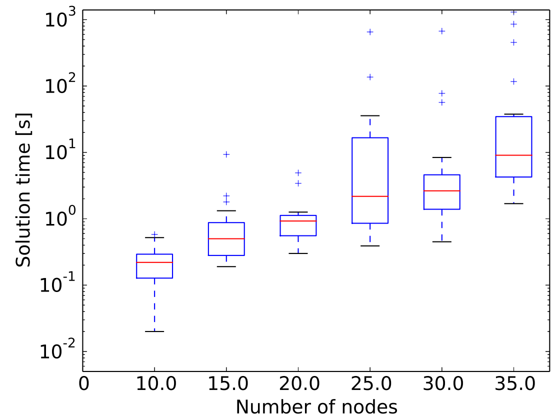


Fig. 19: Solution time for $1K$ routing with min-max energy optimisation, shown in logarithmic scale. The red lines indicate the median, and the blue boxes give the first and third quartiles.

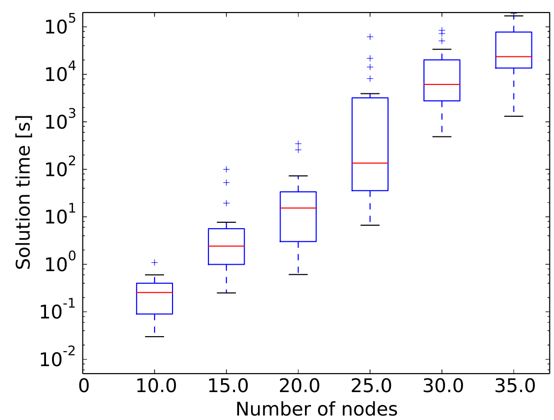


Fig. 20: Solution time for $nK$ routing with min-max energy optimisation, shown in logarithmic scale. The red lines indicate the median, and the blue boxes give the first and third quartiles.

the $1K$ case, and Figure 20 for the $nK$ case. The min-max experiments were executed on an Intel Xeon E5-2420 v2 CPU (2.20 GHz), with 12 virtual cores (6 cores with 2 threads each). Overall, the times to solve problem instances of the same size are similar for min-max energy optimisation and total energy optimisation. Since multiple iterations of min-max optimisation would be required in order to fully characterise the network lifetime, this indicates that optimising for network lifetime is likely to be less applicable in practice, especially for dynamic networks where the routing paths need to be updated frequently. Frame minimisation times, shown in Table V, are somewhat longer for min-max optimisation than for total energy optimisation, but are nonetheless fairly short — on the order of seconds— even for large problem instances. Solving for the reverse arborescences used for routing once again constitutes the bulk of the work.

## IX. Future Work and Conclusions

While the numerical results presented here provide an indication of the performance of our models, there is much more

| N | Frame Length | Solution Time [s] | Iterations | C-sets Generated |
|---|---|---|---|---|
| 10 | 3.0 (±0.24) | 0.02 (±0.01) | 5.7 (±0.54) | 10.85 (±0.78) |
| 15 | 3.15 (±0.25) | 0.09 (±0.02) | 9.2 (±0.74) | 18.4 (±0.81) |
| 20 | 3.65 (±0.47) | 0.17 (±0.04) | 11.55 (±1.9) | 23.5 (±2.03) |
| 25 | 3.6 (±0.45) | 0.35 (±0.11) | 13.6 (±2.83) | 29.55 (±3.38) |
| 30 | 3.95 (±0.55) | 0.54 (±0.12) | 16.25 (±3.28) | 34.65 (±3.47) |
| 35 | 3.55 (±0.32) | 2.82 (±0.84) | 17.8 (±3.41) | 40.6 (±3.57) |

TABLE V: Frame minimisation results for $1K$ scenarios with min-max energy optimisation. Values shown are means across the 20 generated scenarios for each number of nodes, with 95% confidence intervals in parentheses. N = number of nodes.

that remains to be investigated. Firstly, the frame minimisation that was performed here for the $1K$ case should be extended to the $nK$ case. Further, the effects of the different network parameters should be explored more thoroughly. These include the number of destination and origin nodes, the network density, the number of required packets $K$, and the ratio of the aggregation cost to the transmission cost. There is also work to be done in order to better understand the variance observed in the solution times, and whether it can be analysed and predicted from network characteristics and topology. This would aid in taking decisions as to whether an optimal solution is feasible for a given network and time constraint, or whether an approximation needs to be used. The development of approximation algorithms and practical protocols for the problems we have investigated in this work is also an important direction for future work.

There are also a number of possibilities for further development of our optimisation models. First, we have considered two models for the capabilities of the sensor nodes: that they can both aggregate and transit other nodes' measurements, and that they only generate their own measurements, but do not transit or aggregate. However, it may also be the case that some nodes can transit measurements, but not aggregate. This requires substantial changes to the optimisation models and has not been considered in this work.

We have used the total energy usage and min-max energy usage as our objective functions, however, more work is needed in order to fully model network lifetime. Fairness of energy usage between nodes may also be of interest. For latency-constrained applications, energy usage may not be the most important consideration at all, but rather the optimisation should be performed to minimise delay or age of information of sensor readings. Another pressing issue in Internet of Things systems is reliability, which we have not considered here. Greater reliability may be facilitated by, for instance, collecting redundant sensor measurements at multiple destination nodes, or by requiring measurements to be transmitted over multiple, redundant paths.

Emerging Internet of Things networks will rely heavily on machine-to-machine communications. In these networks, fog nodes will provide multiple data collection gateways at the network edge, and actuators will directly utilise sensor measurements. In this work, we have formulated mixed-integer programming problems to optimise routing for both data aggregation and multicast dissemination in these scenarios, as well as throughput optimisation for cases with only unicast aggregation. We have also provided numerical results demonstrating how the performance is affected by the network size. Our models presented here have the potential to provide energy efficient solutions for future networks, as well as performance bounds for approximate algorithms for time- or processing-constrained use cases.

## Acknowledgements

The work of the Polish authors was supported by the National Science Centre, Poland, under the grant no. 2017/25/B/ST7/02313, "Packet routing and transmission scheduling optimization in multi-hop wireless networks with multicast traffic".

## References


[1] F.-M. Tseng, C.-H. Lin, and K.-C. Chen, "In-network computations of machine-to-machine communications for wireless robotics," *Wireless Personal Communications*, vol. 70, no. 3, 2013.

[2] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Workshop on Mobile Cloud Computing MCC*. ACM, 2012, pp. 13–16.

[3] S. Yi, C. Li, and Q. Li, "A survey of fog computing: concepts, applications and issues," in *Workshop on Mobile Big Data*. ACM, 2015, pp. 37–42.

[4] P. Björklund, P. Värbrand, and D. Yuan, "Resource Optimization of Spatial TDMA in Ad hoc Radio Networks: a Column Generation Approach," in *IEEE INFOCOM*, March 2003, pp. 818–824.

[5] A. Capone, G. Carello, I. Filippini, S. Gualandi, and F. Malucelli, "Routing, Scheduling and Channel Assignment in Wireless Mesh Networks: Optimization Models and Algorithms," *Ad Hoc Networks*, vol. 8, no. 6, pp. 545–563, 2010.

[6] P. H. Pathak and R. Dutta, "A Survey of Network Design Problems and Joint Design Approaches in Wireless Mesh Networks," *IEEE Communications Surveys Tutorials*, vol. 13, no. 3, pp. 396–428, 2011.

[7] P. Wan, O. Frieder, X. Jia, F. Yao, X. Xu, and S. Tang, "Wireless Link Scheduling under Physical Interference Model," in *IEEE INFOCOM*, April 2011, pp. 838–845.

[8] M. Pióro, M. Żotkiewicz, B. Staehle, D. Staehle, and D. Yuan, "On Max-min Fair Flow Optimization in Wireless Mesh Networks," *Ad Hoc Networks*, vol. 13, no. 0, pp. 134–152, 2014.

[9] Y. Li, A. Capone, and D. Yuan, "On End-to-end Delay Minimization in Wireless Network under Physical Interference Model," in *IEEE INFOCOM*, April 2015, pp. 2020–2028.

[10] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 14, no. 2, 2007.

[11] J. N. Al-Karaki, R. Ul-Mustafa, and A. E. Kamal, "Data aggregation in wireless sensor networks — exact and approximate algorithms," in *Workshop on High Performance Switching and Routing*. IEEE, 2004, pp. 241–245.

[12] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum lifetime data gathering and aggregation in wireless sensor networks," in *IEEE International Conference on Networking*, 2002, pp. 685–696.

[13] L. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *International Conference on Distributed Computing Systems Workshops*. IEEE, 2002, pp. 575–578.

[14] B. Krishnamachari, D. Estrin, and S. Wicker, "Modelling data-centric routing in wireless sensor networks," in *IEEE INFOCOM*, vol. 2, 2002, pp. 39–44.

[15] D. Li, J. Cao, M. Liu, and Y. Zheng, "Construction of optimal data aggregation trees for wireless sensor networks," in *International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 2006, pp. 475–480.

[16] H. Li, C. Wu, Q.-S. Hua, and F. C. Lau, "Latency-minimizing data aggregation in wireless sensor networks under physical interference model," *Ad Hoc Networks*, vol. 12, pp. 52–68, 2014.

[17] H.-L. Fu, H.-C. Chen, P. Lin, and Y. Fang, "Energy-efficient reporting mechanisms for multi-type real-time monitoring in machine-to-machine communications networks," in *IEEE INFOCOM*. IEEE, 2012, pp. 136–144.

[18] J. L. Wong, R. Jafari, and M. Potkonjak, "Gateway placement for latency and energy efficient data aggregation [wireless sensor networks]," in *IEEE International Conference on Local Computer Networks*. IEEE, 2004, pp. 490–497.

[19] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on Networking (ToN)*, vol. 11, no. 1, pp. 2–16, 2003.

[20] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924–935, 2002.

[21] H. S. Kim, T. F. Abdelzaher, and W. H. Kwon, "Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks," in *International Conference on Embedded Networked Sensor Systems*. ACM, 2003, pp. 193–204.

[22] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, "Impact of network density on data aggregation in wireless sensor networks," in *International Conference on Distributed Computing Systems*. IEEE, 2002, pp. 457–458.

[23] H. Ö. Tan and I. Körpeolu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM Sigmod Record*, vol. 32, no. 4, pp. 66–71, 2003.

[24] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *International Conference on System Sciences*. IEEE, 2000, pp. 10–pp.

[25] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[26] L. Guo, Y. Li, and Z. Cai, "Minimum-latency aggregation scheduling in wireless sensor network," *Journal of Combinatorial Optimization*, vol. 31, no. 1, pp. 279–310, 2016.

[27] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *ACM Sigmod Record*, vol. 31, no. 3, pp. 9–18, 2002.

[28] Y. Yao, J. Gehrke *et al.*, "Query processing in sensor networks." in *Cidr*, 2003, pp. 233–244.

[29] A. Sinha and D. K. Lobiyal, "Performance evaluation of data aggregation for cluster-based wireless sensor network," *Human-Centric Computing and Information Sciences*, vol. 3, no. 1, p. 13, 2013.

[30] V. Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: communication, clustering and aggregation," *Ad Hoc Networks*, vol. 2, no. 1, pp. 45–63, 2004.

[31] M. Handy, M. Haase, and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," in *International Workshop on Mobile and Wireless Communications Networks*. IEEE, 2002, pp. 368–372.

[32] H. Yang and B. Sikdar, "Optimal cluster head selection in the LEACH architecture," in *IEEE International Performance, Computing, and Communications Conference (IPCCC)*. IEEE, 2007, pp. 93–100.

[33] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[34] T. Pham, E. J. Kim, and M. Moh, "On data aggregation quality and energy efficiency of wireless sensor network protocols-extended summary," in *International Conference on Broadband Networks*. IEEE, 2004, pp. 730–732.

[35] H. Çam, S. Özdemir, P. Nair, D. Muthuavinashiappan, and H. O. Sanli, "Energy-efficient secure pattern based data aggregation for wireless sensor networks," *Computer Communications*, vol. 29, no. 4, pp. 446–455, 2006.

[36] K.-W. Fan, S. Liu, and P. Sinha, "Structure-free data aggregation in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 8, 2007.

[37] C.-M. Chao and T.-Y. Hsiao, "Design of structure-free and energy-balanced data aggregation in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 37, pp. 229–239, 2014.

[38] D. Niyato, L. Xiao, and P. Wang, "Machine-to-machine communications for home energy management system in smart grid," *IEEE Communications Magazine*, vol. 49, no. 4, 2011.

[39] I. Stojmenovic, "Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 122–128, 2014.

[40] Y. Li, M. Pióro, D. Yuan, and J. Su, "Optimizing link rate assignment and transmission scheduling in WMN through compatible set generation," *Telecommunications Systems*, vol. 61, no. 2, pp. 325–335, 2016.

[41] M. Pióro, "Network optimization techniques," in *Mathematical Foundations for Signal Processing, Communications, and Networking*, E. Serpedin, T. Chen, and D. Rajan, Eds. Boca Raton, USA: CRC Press, 2012, ch. 18, pp. 627–690.

[42] A. Tomaszewski and M. Pióro, "Packet routing and frame length optimization in wireless mesh networks with multicast communications," in *Proc. of the 17th International Network Strategy and Planning Symposium (NETWORKS 2016), Montreal, Canada*, September 26-28 2016, pp. 3–8.

[43] E. Fitzgerald, "Wireless network scenario generator," https://bitbucket.org/EIT_networking/network_generator.

[44] "NetworkX," https://networkx.github.io, accessed 2017-11-06.