TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES
Trans. Emerging Tel. Tech. 2017; **00**:1–14
DOI: 10.1002/ett

RESEARCH ARTICLE

Energy Efficient Multi-Group Communication

André Riker, Carlos M. Fonseca, Marilia Curado, Edmundo Monteiro

Centre for Informatics and Systems, Department of Informatics Engineering, University of Coimbra, Coimbra, Portugal

ABSTRACT

Wireless devices are widely used to monitor and control multiple groups within the context of Machine-to-Machine applications. The Constrained Application Protocol provides communication capabilities for applications that demand periodic monitoring of multiple groups. Due to the energy constraints of the devices used, a key challenge is to extend the network lifetime. Data aggregation solutions have been proposed to reduce the amount of network traffic and increase energy efficiency. However, for periodic monitoring of multiple groups, current data aggregation solutions do not exploit the potential of combining multiple payloads in a single message. In addition, solutions in literature are unable to take advantage of the communication interactions that occur when there is traffic originating from different groups. To fill this gap, this paper focuses on a non-traditional data aggregation approach, named Two Tier Aggregation, that applies the idea of inserting many payloads in one message to efficiently gather data from multiple groups, introducing novelty on how the messages are assembled. An Integer Linear Programming model is proposed to maximize the network lifetime in multiple group scenarios. The proposed formulation guarantees the energy efficiency of Two Tier Aggregation and defines an upper bound for the heuristics. The evaluation shows the lifetime upper bound obtained by the proposed Integer Linear Programming model on different network sizes, and also compares it to state-of-the-art heuristic solutions.

* **Correspondence**
André Riker, Department of Informatics Engineering, Coimbra, Portugal
E-mail: ariker@dei.uc.pt

1. INTRODUCTION

The Constrained Application Protocol (CoAP) [1] is a web protocol designed for devices with energy, memory, and processing constraints [2]. CoAP has enabled several Machine-to-Machine (M2M) applications, such as smart cities, smart metering, and environmental monitoring [3]. According to the CoAP standards, namely Group Communication [4] and Observing Resources [5], CoAP clients can define multiple groups of nodes, called Monitoring Groups, to *observe* Points-of-Interest, specifying the client's preferred communication settings (e.g. communication periodicity). Thus, it is an important challenge for CoAP applications to improve the energy efficiency of the periodic multi-group communication [6].

Since the early wireless sensor networks, data aggregation has been used as an approach to provide energy efficient communication, because it applies mathematical functions to summarize the data as it is collected and transmitted on each hop. However, the energy benefits provided by data aggregation approaches come with the downside of reducing the information accuracy. Using data aggregation, the clients no longer know the individual data readings produced by each device. Instead, they receive fewer messages containing the aggregated information [7].

Machine-to-Machine networks, which have a large number of devices [8], are more affected by the reduced level of accuracy than small wireless sensor networks. The reason for this is that there are, typically, more divergent data readings in large networks. If the data produced by the network is fully aggregated, the clients will not be able to identify the wide range of measurements within the network. Therefore, the need to control the degree of accuracy of data aggregation in large networks is another aspect that strengthens the idea of using multiple Monitoring Groups to gather data for M2M applications.

Figure 1 illustrates two periodic multi-group applications that use data aggregation, namely Smart Parking and City Monitoring [15]. In the case of Smart Parking, the devices are divided into two groups and are responsible for giving information about the total number of available parking lots every 30 seconds. In the City Monitoring application, there are two groups of devices that can communicate different measurements, such as the levels of carbon dioxide emissions and electricity consumption. Each group communicates and produces an aggregated message revealing the average of the measurements from each group every 60 seconds.

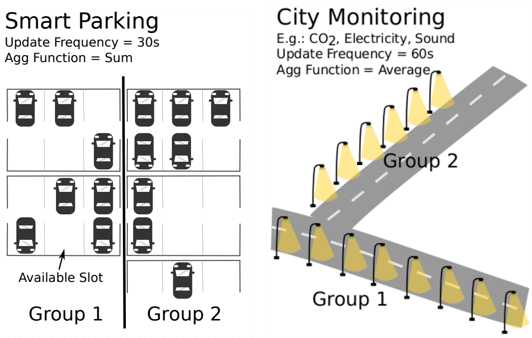


Figure 1. Multi-Group Monitoring Applications.

Considering the context, a relevant M2M research challenge for periodic multi-group communication is to provide energy efficient data aggregation solutions [16]. However, most of the data aggregation solutions do not consider multiple groups nor messages assembled with multiple payloads. Thus, these solutions are unable to exploit the communication interactions that occur when there is traffic originating from different groups. One of the few solutions, proposed by Riker et. al [14], for multi-group scenarios, called Two Tier Aggregation for Multi-target Applications (TTAMA), presents the concept of assembling messages with many payloads. However, the TTAMA's algorithm used to compute the aggregation routes prevents the achievement of the maximum network lifetime, since it uses a static linear function to estimate the cost of the paths.

To fill this gap, the proposed solution adopts a non-traditional data aggregation approach, named Two Tier Aggregation (TTA), that inserts many payloads in one message to efficiently gather data from multiple groups. The contributions of this paper are:

1. It optimizes TTA by formulating a lexicographic multi-objective Integer Linear Program (ILP) that is able to find the network flows that achieve the upper bound of the network lifetime. This formulation is proposed considering single and multiple sinks on the network.
2. It compares the upper bound of the network lifetime with different state-of-the-art heuristic solutions. It conducts a performance evaluation which shows the upper bound of the network lifetime obtained by the proposed solution in networks with different sizes.
3. This paper identifies the main cases where the current heuristics have low network lifetime performance.

Although group-communication within the context of the CoAP protocol represents a strong motivation for the proposed approach, it is not restricted to a specific set of protocols.

The rest of this article is structured as follows: Section 2 shows the main related work. Section 3 describes the Two Tier Aggregation approach and Section 4 outlines the network notation. Section 5 and Section 6 present the proposed solution and the obtained results, respectively. Section 7 discusses the conclusions and future works.

2. RELATED WORK

In Wireless Sensor Networks (WSN), the devices collect data for a single application. M2M changes the WSN requirements by demanding data collection of multiple groups, each one possibly communicating a different data-type, to deliver it for multiple applications. In the M2M context, some recent works have proposed solutions for group-communication. Younghwan et. al [17] propose a group-based communication method to reduce the number of devices accessing the cellular network. Choi et. al [18] also propose a solution for M2M group-communication, but it seeks to improve the security of the M2M groups, which are composed by mobile and static low-power devices. Group-communication solutions designed for M2M capillary networks are relevant for M2M systems. According to Lo et. al [19], M2M capillary networks provide low-power and short-range communication for M2M systems and comprise an important part of most M2M network architectures. Besides, Weyrich et. al [20] strengthens the idea that M2M communication is not restricted to cellular communication, but also involves low-power and short-range protocols such as CoAP and IEEE 802.15.4.

The idea of inserting many payloads inside the same message is also proposed by Stasi et. al [9] and Tsitsipis et. al [10]. These solutions show that a network can significantly improve the communication efficiency by transmitting messages with several payloads combined.

Although these solutions are relevant contributions, they do not address M2M periodic multi-group communication. While Tsitsipis et al. [10] present the idea of messages with multiple payloads as a preliminary concept to ensure high data accuracy and low energy consumption, Stasi et. al [9] focus on using payload concatenation to improve the delay and throughput of multi-path mesh networks.

Another solution that addresses similar aspects is proposed by Bicakci and Tavli [11]. The authors investigate communication strategies for prolonging the network lifetime in multi-domain wireless sensor networks through Linear Programming (LP). The authors consider a scenario of cooperative multi-domain networks deployed in the same physical location, where a different authority manages each domain. The idea of a cooperative multi-domain network is similar to the concept of multi-group networks. The divergent point between our work and Bicakci and Tavli [11] is that the former considers data aggregation procedures to be applied by each group, while the latter does not apply any data aggregation approach on the network traffic.

Among the works that address the problem of maximizing the network lifetime using data aggregation, Kalpakis et al. [12] formulate a LP model to carry this out in single Point-of-Interest scenarios. This LP model is one of the most important references for maximizing the network lifetime by means of Payload Aggregation (e.g. the Maximum, Minimum, Average, and Sum Total), but it does not consider either multiple groups or Payload Concatenation. Hua et al. [13] generalize the formulation of Kalpakis et al. [12], addressing the problem of maximizing the network lifetime and considering other aggregation operations rather than simple mathematical functions. However, they were only concerned with finding efficient heuristics for single-group monitoring applications.

The problem addressed in this paper is different from those described so far, because the TTA approach requires the data producers to be identified to determine which Point-of-Interest a particular message is related to. In the aforementioned related work, the aggregation can be carried out without knowing the data producer, which means the LP formulation is based on the single-commodity network flow problem. In the case of TTA, the aggregation is performed by considering each group as a different data producer, and thus the problem of determining the maximum network lifetime is addressed on the basis of the multi-commodity network flow problem. As each group is a unique data producer, it must have constraints to ensure the appropriate delivery of its data.

Among the works that have addressed the TTA approach, Riker et al. have proposed Data Aggregation

Table I.
Related Work.

Works	Payload Aggregation	Payload Concatenation	Multiple Groups	Network Lifetime Maximization
Stasi et. al [9]	No	Yes	No	No
Tsitsipis et. al [10]	No	Yes	No	No
Bicakci and Tavli [11]	No	No	Yes	Yes
Kalpakis et al. [12]	Yes	No	No	Yes
Hua et. al [13]	Yes	No	No	No
Riker et. al [14]	Yes	Yes	Yes	No

DOI: 10.1002/ett

for Multiple Groups (DAMiG) [21] and Two Tier Aggregation for Multi-target Applications (TTAMA) [14]. Both solutions are heuristics that use cost/weight functions to find the best aggregation paths when using TTA. These solutions show significant improvements in terms of network lifetime when compared with those that do not perform payload concatenation, such as Energy Efficient Spanning tRee (EESR) [22]. Although DAMiG and TTAMA consider the aspects of periodic multi-group communication and adopt the TTA approach to reduce network traffic, these solutions lack a mathematical formulation to determine the maximum network lifetime.

Summing up, as Table I shows, current data aggregation approaches are not propose solutions able to maximize the network lifetime in periodic multi-group scenarios using Payload Aggregation and Concatenation.

3. TWO TIER DATA AGGREGATION APPROACH

Data aggregation can be performed in several ways. In this paper, two important data aggregation techniques are employed:

–**Payload Aggregation:** applies statistical functions to aggregate the payload information into a single value.

–**Payload Concatenation:** produces a message that has a single header and multiple attached payloads, which means that the size of the messages is variable.

A distinction should also be made between Internal and External Group Traffic. When a particular group is taken as a reference-point, *Internal Group Traffic* is the set of messages originating from its members, while *External Group Traffic* is the set of messages produced by other groups.

After these data aggregation techniques have been shown and how to differentiate the traffic from groups, Two Tier Aggregation (TTA) can be defined as a data aggregation approach that applies Payload Aggregation to Internal Group Traffic and Payload Concatenation in External Group Traffic.

Two Tier Aggregation applies the aggregation technique according to the data producer group, using two rules: (i) When the data is in transit inside its producer-group, it is stored in the message as the primary payload and will be aggregated by means of simple mathematical calculations (i.e. the Payload Aggregation technique), such as Maximum, Minimum, Sum, and Average; (ii) If the data goes outside the producer group, it is maintained intact as a secondary payload, which is attached to the primary payload (i.e. the Payload Concatenation technique). Thus, the messages produced by the proposed data aggregation approach have a primary payload and might have multiple secondary payloads attached.

By using the Payload Aggregation technique, TTA can avoid the problem of similar or even repeated messages being sent by each group member. In addition, by means of the Payload Concatenation technique, TTA eliminates redundant headers from the External Group Traffic, and produces messages with a single header and multiple payloads. This saves energy resources that would otherwise be spent on message headers, without impairing the accuracy of the communication, since a node can

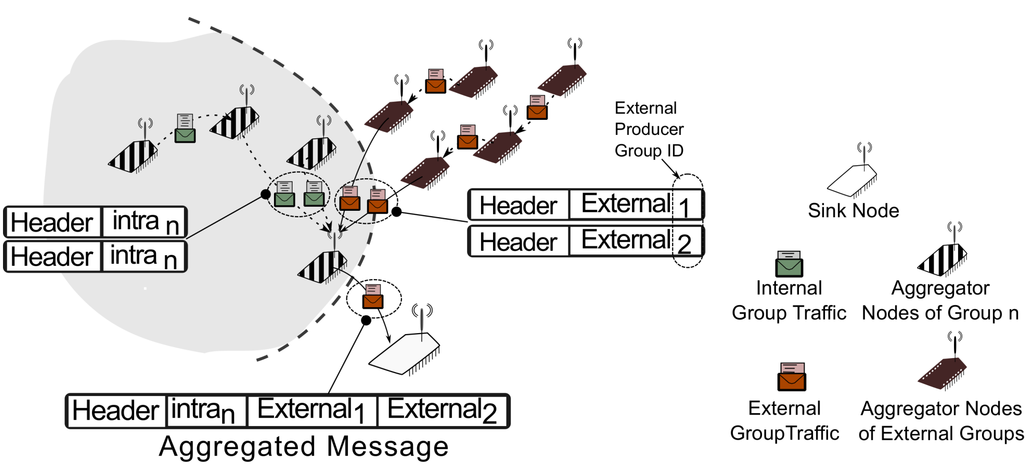


Figure 2. Two Tier Aggregation.

Trans. Emerging Tel. Tech. 2017; **00**:1–14 © 2017 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

suppress the headers of messages produced by another group, and thus keep the payload content intact.

Figure 2 illustrates a particular communication event when a node, which belongs to group n, aggregates Internal and External Group Traffic, by applying the payload aggregation to the Internal Group Traffic and concatenating the payloads derived from the external groups. As can be observed, the produced aggregated message has a single *intra* payload resulting from the payload aggregation operations (e.g. Average, Sum Total, Minimum, and Maximum), but every *external* payload is preserved. Thus, since it is able to employ both techniques, TTA is a suitable data aggregation approach for CoAP messages or other protocol with a similar purpose used to observe multi-groups in the network.

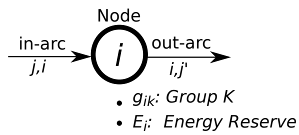
Payload aggregation is only performed on internal group traffic because each group must have a particular communication setting, such as a data aggregation function, communication periodicity, and data types. For instance, group A can be defined by the client to communicate the maximum temperature of a room, while group B gives information about the average amount of light in an office. Thus, it is not possible to compute payload aggregation of groups A and B due to the different settings (i.e. aggregation functions) and the data types. Another reason to perform payload aggregation only on internal group traffic is that the nodes of a particular group are usually located in the same geographic area, tending to present high redundancies on their data. On the contrary, groups located at different areas are not likely to present such high data redundancy.

4. NETWORK MODEL AND PROBLEM STATEMENT

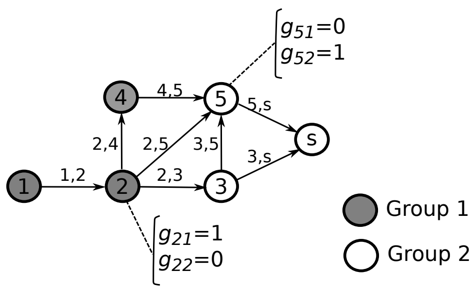
The network is modeled as a directed graph $G(V, A)$, where $V = \{1, 2, ..., n\}$ is the set of nodes and $A = \{(i, j)|i, j \in V\}$ is the set of arcs. The network has a single sink node, denoted by $s \in V$, which is the final destination of all messages. For simplicity, the set of regular nodes (i.e. non-sink nodes) is denoted as $N = V - \{s\}$. In addition, each node i has an energy reserve denoted by E_i, and the monitoring groups are denoted by $S = \{1, 2, ..., k\}$. The following binary values indicate the group that a node i belongs to:

$$g_{ik} = \begin{cases} 1 & \text{if node } i \text{ belongs to group } k \\ 0 & \text{otherwise} \end{cases}$$

Figure 3 illustrates the graph model used. Figure 3.a shows the arcs of a node i, and Figure 3.b presents an example of a graph with 5 nodes and a sink. As can be observed, nodes *1, 2, 4* belong to group 1, while nodes *3* and *5* belong to group 2.



a) Node i of the Graph.



b) Example.

Figure 3. Ilustration of the Graph Model.

The following assumptions are considered in this work:

Assumption 1. *Overlapping of Groups* - A node belongs to a single group. Thus, there is no overlapping of groups.

Assumption 2. *Conservation of Payloads* - The payloads produced in a particular group are conserved by the external members. To illustrate this assumption, suppose the payloads P1 and P2 originating from the same group, A, flow through a particular node external to A. According to this assumption, payload aggregation cannot be used on P1 and P2 to create a single payload.

Besides, the two definitions that bind together communication and network lifetime are given as follows:

Definition 1. *Communication Round* - A particular *communication round* is successfully completed if every node $i \in N$ is able to produce a single payload, and it is delivered to the sink.

DOI: 10.1002/ett

Definition 2. *Node Lifetime* - The *node lifetime* of a node $i \in N$, given by L_i, denotes the total number of communication rounds that can be completed by i considering its energy reserve E_i. Consequently, L_i measures the number of payloads that will be produced and communicated by i before its energy reserve is depleted.

Definition 3. *Network Lifetime* - According to Zhang et al [23], the *network lifetime*, denoted as *NL*, ends when at least one node is no longer able to produce or deliver the produced payload to the sink.

According to Definition 3, $NL = min_{\forall i \in N} L_i$. So, the problem of Maximizing the Lifetime using Two Tier Aggregation (Max-Lifetime TTA) can be stated as follows.

The problem, denominated *Max-Lifetime TTA Problem*, is solved by the set of network flows that achieves the maximum network lifetime and does not violate the rules of payload aggregation and concatenation defined by the Two Tier Aggregation approach.

This problem can be solved by maximizing the minimum L_i. Given that all produced traffic has to be communicated over the network, the solution for the *Max-Lifetime TTA Problem* is the network flow that maximizes the payload production (i.e. NL).

5. NETWORK LIFETIME OPTIMIZATION

The optimization of the Max-Lifetime TTA problem via Integer Linear Programming is presented for single sink scenarios in Section 5.1. Following this, Section 5.2 introduces the changes in the model to maximize the lifetime in the multiple sink scenario.

5.1. Integer Linear Programming Formulation

The number of k-group payloads travelling over the arc *(i,j)* is represented by p_{ijk}. Besides, as it is not necessary to identify the headers in terms of groups, the number of headers travelling over an arc *(i,j)* is denoted by h_{ij}.

Setting a default size for payloads and headers, which will be further detailed in Section 6, it is possible to define *Htx* and *Ptx* as the transmission energy for a node i to transmit a single header and a single payload, respectively. Similarly, *Hrx* and *Prx* can be defined as the energy needed for node i to receive a single header and a single payload on arc *(i,j)*, respectively. Table II summarizes the symbols related to the optimization model.

Table II. Definition of terms.

Symbol	Definition
E_i	Initial energy of node $i \in N$.
p_{ijk}	Number of k-group payloads travelling over the arc (i, j).
h_{ij}	Number of headers travelling over an arc (i, j).
Htx	Energy consumed to transmit a single header.
Ptx	Energy consumed to transmit a single payload.
Hrx	Energy consumed to receive a single header.
Prx	Energy consumed to receive a single payload.
Tx_{setup}	Energy spent related to the CSMA and CCA procedures.
E_{slp}	Energy consumed per communication round in sleeping state.
E_{CPU}	Energy spent by the CPU to perform aggregation on a single payload.

Eq. 1 defines the energy consumed by a node $i \in N$ regarding the communication of data. It is important to notice that Eq. 1 also includes Tx_{setup} as the energy consumed to setup wireless transmissions. Tx_{setup} captures the energy consumption involved in the tasks before the transmission, which include the energy to perform Carrier Sense Multiple Access (CSMA) back off, Channel Clear Assessment (CCA) detection, and the change from Rx (CCA) to Tx mode.

$$Ecomm_i = \sum_{j:(i,j)\in A} h_{ij}(Tx_{setup} + Htx) + \sum_{j:(i,j)\in A} \sum_{k\in S} p_{ijk} Ptx + \sum_{j:(j,i)\in A} h_{ji} Hrx + \sum_{j:(j,i)\in A} \sum_{k\in S} p_{jik} Prx \quad (1)$$

A node also consumes energy performing CPU operations to aggregate data, as presented in Eq. 2.

$$Eagg_i = \sum_{j:(j,i)\in A} \sum_{k\in S} p_{jik} g_{ik} E_{CPU} \quad (2)$$

Header aggregation is not included in Eq. 2, since all headers received by a node i are dropped. The total energy spent of node i on sleeping mode is $Esleep_i$, and it is computed as follows:

DOI: 10.1002/ett

$$Esleep_i = NL\, E_{slp} \tag{3}$$

Here, E_{slp} is a term that represents the amount of energy per communication round spent on keeping the CPU in low power mode and the transceiver is turned off. In Eq. 3, E_{slp} is multiplied by NL because it is the number of communication rounds.

Finally, the total energy consumption of a node $i \in N$ is defined in Eq. 4.

$$Nspent_i = \overbrace{Ecomm_i}^{\text{Communication}} + \overbrace{Eagg_i}^{\text{Aggregegation}} + \overbrace{Esleep_i}^{\text{Sleep mode}} \tag{4}$$

Section 6 further describes the values assigned to all parameters used in the energy model equations, including E_{CPU}, Tx_{setup}, Htx, Hrx, Ptx, Prx, and E_{slp}.

Knowing the particular energy consumption of each node (given in Eq. 4), ES denotes the energy spent by the network, and is defined in Eq. 5.

$$ES = \sum_{i=1}^{n} Nspent_i \tag{5}$$

Besides, it is necessary for the ILP formulation to define a term, λ, that produces values less than 1 when it is multiplied by ES. Thus, λ is defined as:

$$\lambda = \frac{1}{1 + \sum_{i=1}^{n} E_i} \tag{6}$$

Lifetime-optimal solutions might differ with respect to network energy consumed, given by ES.

Considering two lifetime-optimal solutions, S_1^* and S_2^*, the network energy consumed resulting from S_1^* might be higher than S_2^*. Thus, $\max NL$ does not guarantee that the solutions found are the ones that consume the minimum network energy among all lifetime-optimal solutions.

In order to guarantee both objectives, we formulate this problem using a lexicographic multi-objective approach. The primary objective is to maximize NL, while the secondary objective is to minimize the network energy consumption. The secondary objective can be achieved by setting λES as penalty term in the objective function.

This term will be a secondary objective because it will always be less than 1, while NL is an integer greater than or equal to 1. Therefore, defining the objective as $\max\ NL - \lambda ES$, NL becomes preferred over λES.

Eq. 7 defines the multi-objective function as the network lifetime, to be maximized, penalized by λES, leading to a formulation that minimizes the network energy consumption among all lifetime-optimal configurations. The constraints in Eq. 8 model the fact that the total energy consumed by a node i must be less than or equal to its energy reserve E_i.

The constraints in Eq. 9 are related to the payload aggregation executed on Internal Group Traffic. These constraints ensure that each node i produces, in the course of its lifetime, all payloads belonging to group k (to which node i belongs) that travel over the arcs originating in that node. This means that p_{ijk} is not affected by the number of payloads of the same group arriving from other nodes, since payload aggregation functions produce a single output value, regardless of the number of inputs.

$$\max\ NL - \lambda ES \tag{7}$$

Subject to:

$$Nspent_i \le E_i, \qquad \forall i \in N \tag{8}$$

$$\sum_{j:(i,j)\in A} p_{ijk} = NL, \qquad \forall i \in N, \quad \forall k \in S : g_{ik} = 1 \tag{9}$$

$$h_{ij} \ge p_{ijk}, \qquad \forall (i,j) \in A, \quad \forall k \in S \tag{10}$$

$$\sum_{j:(i,j)\in A} \sum_{k\in S} p_{ijk}(1-g_{ik}) = \sum_{j:(j,i)\in A} \sum_{k\in S} p_{jik}(1-g_{ik}), \qquad \forall i \in N \tag{11}$$

$$\sum_{u:(u,s)\in A} p_{usk} \ge NL, \qquad \forall k \in S \tag{12}$$

$$p_{ijk} = 0, \qquad \forall i,j \in N, \quad \forall k \in S : g_{jk} = 1 \wedge g_{ik} = 0 \tag{13}$$

$$NL \in \mathbb{Z}^+, \tag{14}$$

$$h_{ij}, p_{ijk} \in \mathbb{Z}^+, \qquad \forall (i,j) \in A, \quad \forall k \in S \tag{15}$$

It occurs because each numeric value taken as input of the payload aggregation function is defined as a primitive datatype, so both the input and the output values have a fixed number of bits. For a given set of input numbers defined as 32-bit Float, the payload aggregation function will produce a single 32-bit Float number as result.

The constraints in Eq. 10 state that the number of headers travelling on an arc $(i, j) \in A$, which is given by h_{ij}, must be greater than or equal to the greatest payload of group k travelling on this arc, i.e. p_{ijk}. It is important to note that these constraints do not ensure that the number of headers is minimal. However, due to the energy consumption penalty term in the objective function, the number of headers in the final solution will be as small as possible for the corresponding number of produced payloads.

The constraints in Eq. 11 enforce that each node i transmits to its neighbors all the payloads received by i that were produced in another group (i.e. when $g_{ik} = 0$). Hence, these constraints conserve the payloads produced by external groups, avoiding that non-member nodes perform payload aggregation.

The constraints in Eq. 12 ensure a minimum number of payloads delivered to the sink node. These constraints impose that the number of k-group payloads delivered to the sink is greater than or equal to *NL*, which would correspond to the maximum payload aggregation. Besides, the constraints in Eq. 13 exist to avoid payload loops over the groups. These constraints state that a k-group node j (i.e. $g_{jk} = 1$) cannot receive k-group payloads from a node i that does not belong to k (i.e. $g_{ik} = 0$). Finally, the constraints in Eq. 14 and 15 require that the variables *NL*, p_{ijk} and h_{ij} are all integers greater than or equal to zero.

It is important to notice that the literature classifies a Linear Program as Integer Linear Program when all decision variables must be integers [24]. If some, but not all, variables are restricted to be integer, it is called Mixed Integer Linear Program. In the case of our formulation, all decision variables are integer, therefore it is called Integer Linear Program. In some parts, the proposed formulation is not restricted to be integer (e.g. E_i and $Nspent_i$), but the decision variables continue to be integer.

Figure 4 illustrates an optimal solution obtained from the proposed ILP, considering a network composed of six nodes and one sink. The nodes apply Payload Aggregation on the payloads of the Internal Group Traffic and remove the redundant headers from the messages originating in external groups. The existence of two groups creates three commodities in the network, namely payloads originating in groups 1 and 2, and headers. Nodes 1 and 6 do not communicate with external groups, while nodes 2, 3, 4, and 5 perform external group communication. Taking node 5 as example, it receives 20 payloads originating in group 1 from node 3, produces 88 payloads, communicating a total of 108 payloads and 88 headers. It means that node 5 assembled 20 messages with 2 payloads.

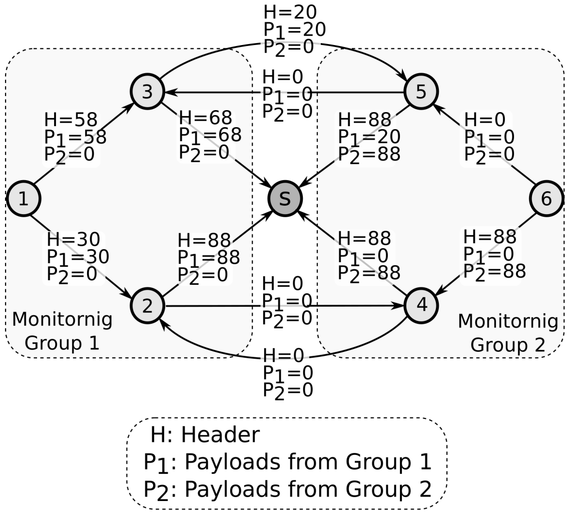


Figure 4. ILP solution example.

5.2. Multiple Sink Formulation

Networks having multiple sinks are an important scenario, especially for large scale monitoring periodic applications. One of the advantages of multiple sinks is that the network has more than one connection with the core network, making the data communication more reliable. Besides, multiple sinks mitigate the communication bottleneck that occurs when there is a single sink over the network.

To optimize the network lifetime using TTA in multiple sinks networks, the optimization model should consider a set of sinks as final destination, which is denoted as $S = \{s_1, s_2, ..., s_w\}$. Thus, a messages going out a particular node has multiple sinks options to be sent, but only one sink is selected as destination, which means that the same message is not received by multiple sinks. Knowing this, the new version of the Eq. 11 should consider all $s_m \in S$, as it is presented in Eq. 16:

Trans. Emerging Tel. Tech. 2017; **00**:1–14
DOI: 10.1002/ett

$$\sum_{s_m \in S} \sum_{u:(u,s_m) \in A} p_{us_m k} \geq NL, \; \forall k \in S \qquad (16)$$

This new constraint states that payloads delivered to all sinks will count as valid payload communication. Special attention should be given to this constraint since it can be easily misunderstood. Eq. 16 ensures that the number of payloads delivered to any sink must be equal or greater than the *NL*. As *NL* = min L_i, the constraints in Eq. 16 ensure that the sinks cannot receive less than the min L_i, since min L_i is the least number of produced payloads. For instance, the constraints in Eq. 16 force the sinks to receive at least one payload from each group k in a communication round *c*. In some cases, depending on the layout and the size of the group k, the sinks will receive more than one payload from group k in a communication round *c*.

This constraint also influences the level of the payload aggregation performed on the Internal Group Traffic. It is worth to observe that Eq. 16 limits the *maximum* aggregation level allowed on the Internal Group Traffic. On the opposite side, the case of minimum aggregation level occurs when the set of sinks receives all payloads produced by the members of the group, which means payload aggregation does not take place.

6. EVALUATION AND RESULTS

This section shows the performance of the proposed ILP model measured in different network and group sizes. Section 6.1 gives details about the values assigned to the parameters of the ILP model. Section 6.2 presents the group and network sizes and layout. Section 6.3 shows the network lifetime and the energy consumption of the proposed ILP, and also presents a comparison between the upper bound obtained by the ILP and the heuristics taken from literature, namely Data Aggregation for Multiple Groups (DAMiG) [21], Two Tier Aggregation for Multi-target Applications (TTAMA) [14], and Energy Efficient Spanning tRee (EESR) [22].

6.1. Parameter Configuration

The values assigned to the parameters of the proposed ILP are based on the ATMega128L [25] micro-controller and the CC2420 transceiver. Based on IEEE 802.15.4 [26], the data rate communication defined as 250 Kbps, and the communication round periodicity is 10s. Most of the non-decision variables are shown in Table III. At 250 Kbps, the time necessary to communicate (i.e. Tx or Rx) a header of 48 Bytes is 0.037ms. Similarly, the time to communicate (i.e. Tx or Rx) a payload of 4 bytes is 0.012ms. The proposed solution is not restricted to a specific technology or protocol. However, reference standards are used to define the values of the parameters. CoAP [1] has been used as reference to define the size of the payload. Regarding the size of the header, we considered a protocol stack composed of CoAP, UDP, 6LowPAN, and IEEE 802.15.4.

According to Casilari et al. [27], before each transmission, a node has to perform setup procedures, which include the periods of Carrier Sense Multiple Access (CSMA) back off (2.24ms), Channel Clear Assessment (CCA) detection (0.128ms), and turnaround from Rx to Tx mode (0.192ms). Thus, Tx_{setup} is 2.56ms x 17.4mA x 3v= 0.134 mJ.

Table III. Values for the Energy Consumption Model.

Symbol	Energy Value	Duration	Current
Htx	0.6438μJ	0.037 ms	17.4 mA
Tx_{setup}	0.134mJ	2.56ms	17.4 mA
Hrx	0.6956μJ	0.037 ms	18.8 mA
E_{CPU}	1.24μJ	51.7 μs	5 mA
Ptx	0.2088μJ	0.012 ms	17.4 mA
E_{slp}	–	–	15 μA
Prx	0.2256μJ	0.012 ms	18.8 mA
E_i	10J	–	–

Default Voltage is 3V.

Regarding the computation costs, each cycle of ATMega128L lasts 1.25e^{-7}s. The Application Report* shows that the ATMega128L executes on average 414 cycles to compute one simple math operation (i.e. addition, subtraction, multiplication, or division) on two 32-bit float numbers. As the aggregation functions considered in this article involve also simple math operations, we consider 414 as the number of cycles necessary to aggregate two data payloads. This number of cycles lasts 1.25e^{-7} x 414 = 51.7μs. Thus, E_{CPU} corresponds to 51.7μs x 5mA x 3v = 1.24 μJ. For all nodes, the energy reserve E_i corresponds to 10 J. Finally, the time in sleeping mode per communication round corresponds to the time in which the node is not in any other state.

* www.ti.com/lit/an/slaa205c/slaa205c.pdf

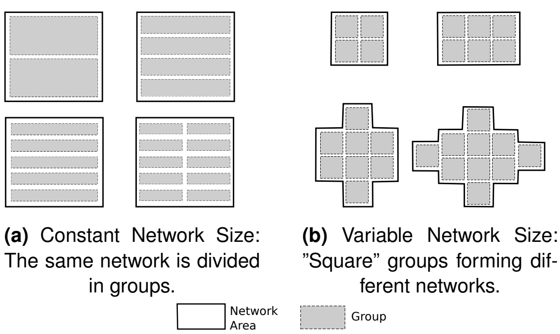
DOI: 10.1002/ett

6.2. Layout and Size of Groups

The size and layout of the network and groups have a strong influence on the energy consumption and on the network lifetime. In order to evaluate a representative numbers of scenarios, the evaluation has followed two approaches regarding the layout, size and number of groups.

- **Constant network size with variable number of group members (Constant Network Size):** In this approach, a network size is defined and the groups with equal number of members are distributed over the network. Seeking the fairness between the groups, all groups should have the same number of members, which means that it is possible to find the group size by finding the natural divisors of the network size. For instance, in the case of a network having 100 nodes, the group sizes that will form groups with the same number of members are: 1, 2, 4, 5, 10, 20, 25, 50, and 100.
- **Variable network sizes with constant number of group members (Variable Network Size):** In this approach, the number of members for each group is constant, while the number of groups and network size are variable. In this case, the fairness between the groups is also kept, since all groups should have the same number of members. For instance, defining the size of the groups as 4 and the total number of groups as 10, the network size will be 40. To obtain a ”square” layout, the groups sizes are given by power of two (e.g. 2x2, 3x3, and 4x4).

Figure 5 illustrates the two approaches applied to form the evaluated networks and groups.



(a) Constant Network Size: The same network is divided in groups.

(b) Variable Network Size: ”Square” groups forming different networks.

Figure 5. Groups Distribution.

All the evaluated nodes have a wireless range of 20m, while the horizontal and vertical spacing between the nodes is 20m, so the nodes have a maximum of 4 neighbors.

6.3. ILP Network Lifetime and Energy Consumption

All results of non-deterministic experiments were collected from 30 runs and the average values were computed. Due to the deterministic nature of some part of the evaluation, the variation of the results showed to be negligible without loss of statistical validity.

The main performance metric used in this evaluation is Network Lifetime (see Def. 3). Besides, as the energy consumption is a secondary objective, it is also used as a performance metric.

Figure 6 shows the network lifetime obtained by the proposed ILP considering a variable network size with constant number of group members (see Figure 5.b). The network lifetime is presented for networks having 4, 6, 8, and 10 groups.

As can be seen, the number of groups over the network impacts the network lifetime, since it controls the amount of External Group Traffic. However, the network lifetime is not necessarily impacted when the number of members of each group increases. For instance, considering a network with 10 groups, the obtained network lifetime is almost constant when the group size is 25, 36, 49, and 64 nodes.

These results (Figure 6) indicate that for a particular number of groups, the ILP model reaches a stable lifetime even if the number of group members increases. The reason for the lifetime stability is that the number of

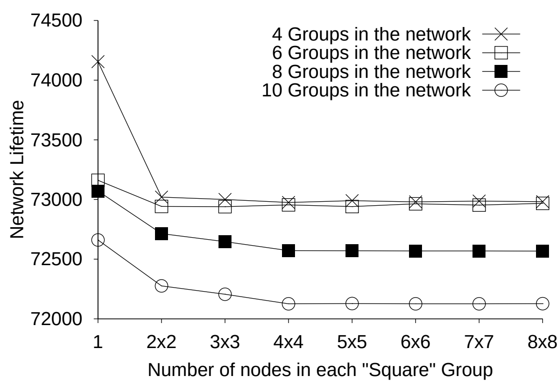


Figure 6. ILP Network Lifetime (Variable Network Size).

Trans. Emerging Tel. Tech. 2017; **00**:1–14
DOI: 10.1002/ett

payloads going outside a group is reduced to one regardless the group size.

Considering a constant network size with variable number of group members (See Figure 5.a), Figure 7 shows the network lifetime and the total network energy consumption due to the transmission of headers, internal group payloads, and external group payloads. There are fewer internal group payloads when the groups become smaller, causing a decrease in energy consumption. Due to External Group Traffic, the energy consumption for 50 monitoring groups is the highest, and the lifetime is the lowest.

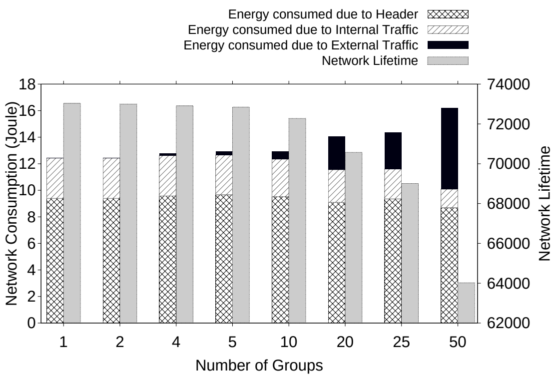


Figure 7. ILP Lifetime and Energy Consumption (Constant Network Size - 100 nodes).

Another relevant aspect is the aggregation behavior of the optimal solution obtained by the ILP. In this direction, Table IV shows for different number of groups the average number of produced, delivered and aggregated payloads.

The percentage of payload aggregation in Table IV is the ratio between the sum of payloads delivered to the sink and the sum of payloads produced. The maximum Internal Group Aggregation that a group can achieve is limited by the number of members. For example, when a network of 100 nodes has 50 groups, the maximum Internal Group Aggregation is 50%, since there are 2 members in each group and at least 1 payload should be communicated to the sink. For the case of 20 groups, each group has 5 members, which means that the maximum Internal Group Aggregation is 4/5= 80%. It is possible to notice that when the network has 1 and 2 groups, the Internal Group Aggregation is around 95%, which is not the maximum Internal Group Aggregation. However, when the network has 4, 5, 10, 20, or 50 groups, the percentage of aggregation decreases, achieving the maximum allowed Internal Group Aggregation. In our evaluation, TTAMA shows a similar level of payload aggregation, but it does not achieve the same results of the ILP when there are many groups. This indicates that TTAMA uses a correct approach to perform payload aggregation, but fails to find the paths to perform efficient payload concatenation.

Table IV. ILP Payload Statistics (Constant Network Size - 100 nodes).

Number of Groups	Payloads Produced	Delivered	Aggregated*	Aggregation
1	7303000	292120	7010880	96 %
2	3651400	146056	3505344	96 %
4	1822650	72910	1749740	96 %
5	1456840	72944	1383896	95 %
10	722740	72184	650556	90 %
20	352850	70492	282358	80 %
50	128040	64466	63574	50 %

*These payloads do not go outside the group.

Another interesting point of evaluation is the comparison between the network lifetime obtained from the ILP and the state-of-art heuristics, namely DAMiG [21], TTAMA [14], and EESR [22]. As Table V presents, the proposed ILP solution has the best performance, which is expected because it achieves the network lifetime upper bound of TTA. The reasons for the lack of performance of the heuristics are: (i) the heuristics are based on link cost functions that find aggregation paths. Besides, these solutions rely on static rules to improve network lifetime via payload and header aggregation. For instance, in DAMiG and EESR, all traffic of a group must pass through a single node before going to an external group. This static rule is not efficient for a small number of groups. Therefore these heuristic solutions cannot achieve the optimal network lifetime; (ii) DAMiG does not achieve a higher lifetime performance because it always seeks to maximize the internal group traffic aggregation, which is not efficient in case of large groups; (iii) EESR does not apply header aggregation, which contributes to the poor lifetime performance.

TTAMA corrects the problem of DAMiG for the case of few groups (i.e. 1 and 2). When the number of groups is very small, the comparison shows that TTAMA achieves a performance that corresponds to more than 90% of the ILP. The main problem of TTAMA is related to the communication of external group traffic. As can be noticed, TTAMA does not maintain the same performance of ILP when there are many groups and the amount of

Table V. ILP and Heuristics Comparison (Constant Network Size - 100 nodes).

	Number of Groups							
	1	**2**	**4**	**5**	**10**	**20**	**25**	**50**
ILP Network Lifetime	73030	73028	72906	72842	72274	70570	69006	64020
TTAMA's Network Lifetime	72190	67030	55770	51160	39830	40060	38520	38204
TTAMA and ILP Ratio	98.85%	91.79%	76.50%	70.23%	55.11%	56.77%	55.82%	59.68%
DAMiG's Network Lifetime	36590	24570	17660	16280	13250	11340	10980	7870
DAMiG and ILP Ratio	50.1%	33.6%	24.22%	22.35%	18.33%	16.07%	15.91%	12.29%
EERS's Network Lifetime	36590	24610	14740	12270	7020	3660	3250	1550
EERS and ILP Ratio	50.1%	33.7%	20.22%	16.84%	9.71%	5.19%	4.71%	2.42%

external group traffic increases. It suggests that the static function used by TTAMA to find the best path to perform payload concatenation is not efficient.

It is well-known that Integer Linear Programming problems belong to the class of Non-deterministic Polynomial (NP) problems. It means that unless P = NP, there is no polynomial-time algorithm for solving ILP problems. Although the complexity of the ILP problems, it is important to highlight that the proposed Integer Linear Program is computed once for the entire network lifetime, and it is used to compute the network lifetime upper bound. Thus, the ILP does not need to be periodically computed, which reduces considerably the amount of resources spent to find the maximum network lifetime.

A computer with 1,8 GHz Intel Core i5 and 4 GB of RAM was used to solve the proposed ILP model. Considering the case where there is a single group on a network of 100 nodes, the time to compute the optimal solution is on average 0.05 seconds. For the case where the network has 100 nodes and the number of groups is 10, the time to find the solution is 1.33 seconds.

Motived by the fact that the proposed ILP demands a considerable amount of computational resources, especially for large network sizes, it is important to present alternatives to solve the proposed optimization model in an affordable amount of time. One technique largely present in the literature to decrease the complexity of the ILP problems is to relax the integrality constraints. With this technique, instead of having constraints forcing the values of the variables to belong to the Integer set $\mathbb{Z}$, the values of the variables are allowed to belong to the Rational set $\mathbb{Q}$. Thus, the Integer Linear Program turns into a Linear Program (LP).

The relaxation of the ILP model raises the question of how large the gap between the optimal (i.e. ILP) and the relaxed-LP model is. In our case, the relaxed-LP model

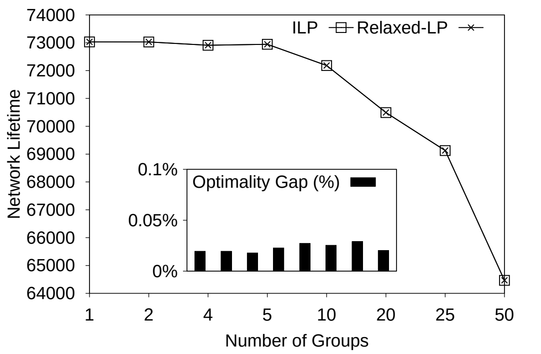


Figure 8. Optimality gap: difference between the relaxed model and ILP.

drops the integrality constraints and has $\lambda = 0$. Figure 8 shows the optimality gap between the two models. As can be seen, the optimality gap is not greater than 0.05 %. This low percentage certifies that the relaxed-LP is not distant from the optimal solution.

For the evaluation of larger network sizes, the relaxed-LP model was used. The network lifetime performance shown in Figure 9 considers networks with up to 4000 nodes, varying the size of the message headers. As can be seen, the header size has great influence on the network lifetime. This influence occurs because the header communication is the most demanding energy

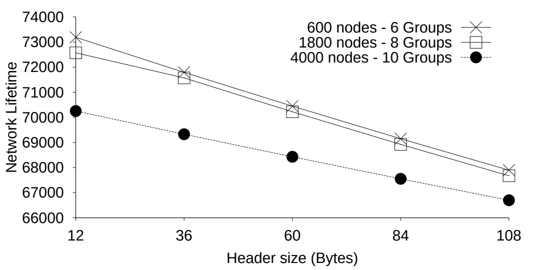


Figure 9. Network lifetime of large scenarios using different header sizes (Constant Network Size).

Trans. Emerging Tel. Tech. 2017; **00**:1–14
DOI: 10.1002/ett

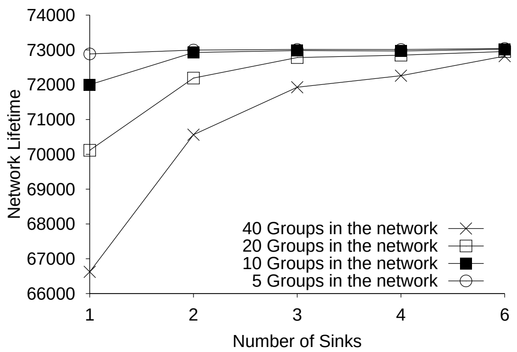


Figure 10. Network lifetime of Relaxed LP with Multiple Sinks (Constant Network Size - 400 Nodes).

activity (see Figure 7), and longer headers reduce the payload size inside each message, which limits the payload concatenation and thus the network lifetime.

Considering network scenarios with multiple sinks, Figure 10 shows the network lifetime performance considering 1, 2, 3, 4, and 6 sinks, while the number of groups is 40, 20, 10, and 5. In this evaluation, the network has 400 nodes. As can be seen, the network lifetime is less influenced by the number of sinks when there are 5 groups. For all other scenarios, the network lifetime presents significant variation. The reason for it is that the nodes have an energy consumption constraint, and the increase on the number of sinks will not change the energy constraint. However, when the number of groups increases, a communication bottleneck is created near the sink. Thus, for the cases of 10, 20, and 40 groups, increasing the number of sinks mitigates the communication bottleneck, which improves the network lifetime performance.

7. CONCLUSION AND FUTURE WORKS

M2M standards have included communication capabilities to address periodic multi-group applications. The data aggregation solutions have to evolve to efficiently support periodic multi-group communication.

This paper studies the network lifetime upper bound of the Two Tier Aggregation (TTA) approach. It proposes and assesses the ILP model in different scenarios and compares it to three state-of-art heuristics. The proposed ILP solution is designed to be solved once for the entire network lifetime and it must be solved by an external device that does not have severe energy constraints like battery-powered sensors.

The obtained results show that the heuristics have presented improvements to support periodic multi-group communication. However, based on the upper bound performance provided by the proposed Integer Linear Program, it is possible to identify lack of performance in some scenarios. As shown in this paper, these problems are related to the inefficiency of using payload concatenation to communicate external group traffic.

As future work, we will implement Two Tier Aggregation (TTA) approach in real-devices. The testbed will allow the proposed optimization to be evaluated considering typical limitations of the devices and also a dynamic environment.

ACKNOWLEDGMENT

This work was partially funded by Ciencia sem Fronteiras (Brazil) number 246599/2012-0 and partially carried out in the scope of the MobiWise project: From mobile sensing to mobility advising (P2020 SAICTPAC/0011/2015), co-financed by COMPETE 2020, Portugal 2020 - POCI, European Unions ERDF, and the Portuguese Foundation for Science and Technology (FCT). This work was also co-financed by the DenseNet project PTDC/EEI-SCR/6453/2014, which is a FCT/FEDER/COMPETE 2020 project.

REFERENCES

1. Shelby Z, Hartke K, Bormann C. The Constrained Application Protocol (CoAP). RFC 7252 jun 2014. URL http://www.ietf.org/rfc/rfc7252.txt.
2. Bormann C, Castellani AP, Shelby Z. CoAP: An application protocol for billions of tiny internet nodes. *IEEE Internet Computing* 2012; **16**(2):62.
3. Kovatsch M. CoAP for the web of things: from tiny resource-constrained devices to the web browser. *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, ACM, 2013; 1495–1504.

4. Rahman A, Dijk E. Group Communication for the Constrained Application Protocol (CoAP). RFC 7390 (Experimental) Oct 2014. URL http://www.ietf.org/rfc/rfc7390.txt.
5. Chunduri U, Tian A, Lu W. The Keying and Authentication for Routing Protocol (KARP) IS-IS Security Analysis. RFC 7645 (Informational) Sep 2015. URL http://www.ietf.org/rfc/rfc7645.txt.
6. Ferdouse L, Anpalagan A, Misra S. Congestion and overload control techniques in massive m2m systems: a survey. *Transactions on Emerging Telecommunications Technologies* 2015; .
7. Li H, Lin K, Li K. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks. *Computer Communications* 2011; **34**(4):591–597.
8. Stojmenovic I. Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems. *IEEE Internet of Things Journal* 2014; **1**(2):122–128.
9. Di Stasi G, Karlsson J, Avallone S, Canonico R, Kassler A, Brunstrom A. Combining multi-path forwarding and packet aggregation for improved network performance in wireless mesh networks. *Computer Networks* 2014; **64**:26–37.
10. Tsitsipis D, Dima SM, Kritikakou A, Panagiotou C, Koubias S. Data merge: A data aggregation technique for wireless sensor networks. *Emerging Technologies & Factory Automation (ETFA), 2011 IEEE 16th Conference on*, IEEE, 2011; 1–4.
11. Bicakci K, Tavli B. Prolonging network lifetime with multi-domain cooperation strategies in wireless sensor networks. *Ad Hoc Networks* 2010; **8**(6):582–596.
12. Kalpakis K, Dasgupta K, Namjoshi P. Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks. *Computer Networks* 2003; **42**(6):697–716.
13. Hua C, Yum TSP. Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks. *IEEE/ACM Transactions on networking* 2008; **16**(4):892–903.
14. Riker A, Cerqueira E, Curado M, Monteiro E. A two-tier adaptive data aggregation approach for M2M group-communication. *IEEE Sensors Journal* 2016; **16**(3):823–835.
15. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M. Internet of things for smart cities. *IEEE Internet of Things Journal* 2014; **1**(1):22–32.
16. Tirronen T, Larmo A, Sachs J, Lindoff B, Wiberg N. Machine-to-machine communication with long-term evolution with reduced device energy consumption. *Transactions on Emerging Telecommunications Technologies* 2013; **24**(4):413–426.
17. Jung Y, Kim D, An S. Scalable group-based machine-to-machine communications in lte-advanced networks. *Wireless Networks* 2013; :1–12.
18. Choi D, Choi HK, Lee SY. A group-based security protocol for machine-type communications in lte-advanced. *Wireless Networks* Feb 2015; **21**(2):405–419.
19. Lo A, Law Y, Jacobsson M. A cellular-centric service architecture for machine-to-machine (m2m) communications. *IEEE Wireless Communications* 2013; **20**(5):143–151.
20. Weyrich M, Schmidt JP, Ebert C. Machine-to-machine communication. *IEEE Software* 2014; **31**(4):19–23.
21. Riker A, Cerqueira E, Curado M, Monteiro E. Data aggregation for group communication in machine-to-machine environments. *Wireless Days (WD), 2014 IFIP*, IEEE, 2014; 1–7.
22. Hussain S, Islam O. An energy efficient spanning tree based multi-hop routing in wireless sensor networks. *Wireless Communications and Networking Conference (WCNC)*, IEEE, 2007; 4383–4388.
23. Zhang X, Anpalagan A, Guo L, Khwaja AS. An energy-aware load-balanced routing protocol for ad hoc m2m communications. *Transactions on Emerging Telecommunications Technologies* 2015; .
24. Bradley S, Hax A, Magnanti T. *Applied mathematical programming*. Addison Wesley, 1977.
25. Atmel. Atmega128, accessed: jan 2016. URL http://www.atmel.com/images/doc2467.pdf.
26. IEEE. Ieee standard for low-rate wireless networks. *IEEE Std 802.15.4-2015 (Revision of IEEE Std 802.15.4-2011)* April 2016; :1–709doi:10.1109/IEEESTD.2016.7460875.
27. Casilari E, Cano-García JM, Campos-Garrido G. Modeling of current consumption in 802.15. 4/zigbee sensor motes. *Sensors* 2010; **10**(6):5443–5468.