Abstract— In this paper, the energy efficiency of edge computing platforms for IoT networks connected to a passive optical network (PON) is investigated. We have developed a mixed integer linear programming (MILP) optimization model, which optimizes the placement and number of the cloudlets and VMs and utilizes energy efficient routes with the objective of minimizing the total IoT network and processing power consumption. Our results show that the power consumption can be reduced by consolidating the placement of these VMs at the PON Optical Line Terminal (OLT) in cases where the traffic volume is still high after data processing, i.e., at low traffic reduction percentages. On the other hand, at high traffic reduction ratios, better power efficiency can be accomplished by placing VMs in lower layer nodes (relays). Our results indicate that utilizing PONs and serving heterogeneous VMs can save up to 19% of the total power. Based on the MILP model insights, a heuristic is developed with very comparable MILP-heuristic power consumption values. We considered three scenarios that represent different levels of homogeneous and heterogeneous VM CPU demands. Good agreement was observed between the heuristic results (17% power saving) and the MILP which results in 19% power saving.

Index Terms—IoT; Passive Optical Networks; Virtual Machines; Edge Computing; Energy Efficiency

I. INTRODUCTION

As a result of the exponential growth of the Internet traffic, the CO₂ emissions and energy consumption of information and communication technology (ICT) networks are undergoing a dramatic increase. This increase is one of the significant challenges that may hinder the expansion of the Internet. Moreover, ICT generates an estimated 2% of the global CO₂ emissions [1]. Consequently, more attention must be given to improving energy efficiency and sustainability of the Internet and the ICT industries.

IoT represents a major evolution in legacy data communication. It is predicted that there will be 75 billion IoT interconnected devices by 2025 [2]. This growing level of connected devices has paved the way for futuristic smart applications in healthcare, agriculture, transportation, manufacturing, smart homes and machine-to-machine (M2M) communications [3], [4]. There are, however, many key challenges such as reliability, security, interoperability and scalability [5]. In addition, one of the main challenges that must be confronted by IoT architects is the energy efficiency and greening the networks [6], which is currently garnering attention in both the academic and the industrial arenas. IoT is also expected to benefit from the wide spectrum of proposed energy efficient network solutions. Cloud computing was investigated as one of the solutions that can improve the utility of IoT by storing and processing the IoT generated data. The energy efficiency of cloud data centres was investigated in [7]-[11]. Virtualization can help improve resource sharing in IoT networks and the supporting data centres and networks, and this was evaluated in [12]-[14]. IoT nodes are typically connected to the access layer of the network [14] and therefore the energy efficiency of this layer as well as that of the metro and the core network have to be improved to improve the overall energy efficiency of the IoT-to-cloud or IoT-to-edge processing architectures. Attention was given to the energy efficiency of different network segments [15]-[21], to the use of renewable energy in these networks to reduce CO₂ [22] and to different energy efficient transmission strategies [23], [24]. IoT nodes can generate large amounts of data and therefore these big data networks have to be optimized to improve their energy efficiency given data processing and networking power consumption, and these were evaluated in [25]-[28]. The exceptional amount of data generated by IoT objects is currently estimated at 2.3 trillion gigabytes of data every day [3].

Serious concerns have been raised about the cost of the energy needed to transport such huge data through the Internet so that it is accessible by anyone anywhere. The connection between the IoT objects and the Internet is facilitated by access networks. One of the most favourable access networks in terms of high bandwidth, long access distance and power consumption is passive optical networks (PON).

Energy constraints are a dominant trait of most IoT end nodes. Many of the IoT implementations use wireless for connectivity. The IoT wireless modules are well known for their hunger for energy. Therefore, processing and computation offloading to the edge of the network is a key method to save energy [3], [29].

Edge computing is proposed to assist in tackling the computational resource poverty of IoT objects. Some of previous studies and research efforts have considered...
addressing some issues such as power consumption, cost and bandwidth in IoT and PON architectures. The authors of [30] proposed a dynamic bandwidth allocation scheme for converged 5G mobile fronthall and IoT networks on TDM-PON. They proposed this scheme to address some technical issues regarding uplink bandwidth management. The work introduced in [31] proposed the implementation of monitoring and control systems in hospice environment through the use of wireless sensors and actuators modules and through the storage of the data in the cloud within a hospital. The authors of [31] proposed the integration of cloud networking with WiFi and ZigBee to realize a Wireless Hospital Digital Interface (WHDI). The authors of [32] improved cost and power consumption figures of the introduced access network by introducing a novel network architecture. The proposed architecture conquers the limitations of both long-reach PONs and mobile backhauling schemes. The enhanced architecture is based on adaptive ultra-long reach links to bypass the Metropolitan Area Network on the core side, in addition to the use of a low cost and low power consumption technology (short-range XPON, wireless) at the end user side. In order to enhance performance, open access networking models and Software Defined Networking (SDN) principles support network virtualization and efficient resource management. In [33], the authors considered some potential IoT access network technologies and examined these technologies over a range of traffic levels in term of power efficiency. The authors of [33] showed that the use of WiFi with PON backhaul, 4G wireless (LTE) access and also GPON access is the most energy efficient access architecture for different IoT traffic levels.

In this paper, we design a framework for an energy efficient edge computing platform for IoT supported by a PON. In this paper, we expand our brief initial work proposed in [34] and provide a full MILP optimization model whose details are given here for the first time. In addition, we expand the work in [34] by developing a heuristic algorithm that mimics, in real time, the behavior of the MILP model introduced.

The remainder of this paper is organized as follows. In Section II, we describe our energy efficient MILP optimization model. Section III discusses the MILP model results, Section IV presents the heuristic and Section V discusses the heuristic results. Finally, in Section VI we give our conclusions.

II. MILP FOR ENERGY EFFICIENT PON-IoT NETWORKS

Our MILP model considers the architecture shown in Figure 1. In this architecture, the upper core network receives the aggregated processed traffic from two separate IoT networks through a PON. In our framework, each IoT network consists of four layers. IoT objects represent the first layer while relay elements represent the second layer. Relay elements aggregate the traffic from the IoT objects. The traffic from the relay elements is aggregated by a single coordinator element in the third layer. A single gateway element is hosted by the last layer in the IoT network. The coordinator traffic is aggregated by this single gateway element and uploaded to the access network (PON). In our framework, the PON access network is made up of two layers. Two ONU entities are hosted by the first layer (ONU layer) while OLT layer consists of a single OLT entity. The task of PON access network is to aggregate the traffic from the IoT network and upload it to the core network.

In our architecture, the entities in PON access network layers and the elements in the three upper layers in each IoT network are allowed to host VMs. These VMs are capable of processing the aggregated traffic. We considered different applications by modelling different types of VMs. Only one VM type is requested by each IoT object. Usefull information is generated by VMs through reducing traffic at different traffic reduction percentages. This traffic reduction is done by the VMs by processing the incoming raw data.

Minimizing the total power consumption is the objective of our proposed MILP. The total power consumption consists of two basic components. Firstly, the power consumption due to traffic through all layers of the proposed architecture. Secondly, the power consumption due to processing by VMS in all possible hosting layers. The MILP power minimization is has to observe some constraints. These constraints consider cloudlet placement, the optimal placement of VMs, managing traffic direction and the traffic flow conservation for unprocessed and processed IoT traffic. Regarding the proposed MILP notations, we have used superscripts to index the variables and the parameters while we have used subscripts as indices of these variables and parameters. Table I defines the parameters used in the MILP model:

| Table I List of parameters and their definitions |
|-------------------------------|-------------|
| Notation | Description |
| $O$ | Set of IoT objects |
| $R$ | Set of relays |
| $C$ | Set of coordinators |
| Variable | Description |
|----------|-------------|
| $G$      | Set of gateways |
| $ONU$    | Set of ONUs |
| $OLT$    | Set of OLTs |
| $TN$     | Set of all IoT network nodes ($TN = O \cup R \cup C \cup G \cup ONU \cup OLT$) |
| $N_x$    | Set of neighbours of node $x$ ($N_x \subseteq TN$) |
| $CN$     | Set of candidate nodes for the cloudlet placement ($CN = R \cup C \cup G \cup ONU \cup OLT$) |
| $VM$     | Set of virtual machine types |
| $\lambda_{ov}^{upt}$ | Un-processed traffic from the IoT object $o$ to the virtual machine $v$ placed at the cloudlet $c$ |
| $\lambda_{oc}^{upt}$ | Un-processed traffic from IoT object $o$ to cloudlet $c$ placed in the candidate networking element |
| $\lambda_{ocxy}^{upt}$ | Un-processed traffic from the IoT object $o$ to cloudlet $c$ placed in the candidate networking element passing through the link between the nodes pair $(x,y)$ |
| $\lambda_{xy}^{upt}$ | Un-processed traffic between the nodes pair $(x,y)$ |
| $\lambda_{cl}$ | Processed traffic from cloudlet $c$ placed in the candidate networking element to the OLT $l$ |
| $\lambda_{clxy}$ | Processed traffic from cloudlet $c$ placed in the candidate networking element to the OLT $l$ passing through the link between the nodes pair $(x,y)$ |
| $I_{vc}$ | $I_{vc} = 1$ if the virtual machine $v$ is placed in the cloudlet $c$, otherwise $I_{vc} = 0$ |
| $H_c$    | $H_c = 1$ if a cloudlet $c$ is built at the candidate networking element, otherwise $H_c = 0$ |
| $TW_c$   | Total normalized workload of the cloudlet $c$ built at candidate networking element |
| $PC^{rp}$ | Total processing induced power consumption of the relays |
| $PC^{cp}$ | Total processing induced power consumption of the coordinators |
| $PC^{gp}$ | Total processing induced power consumption of the gateways |
| $PC^{up}$ | Total processing induced power consumption of the ONUs |
| $PC^{lp}$ | Total processing induced power consumption of the OLTs |
| $PC_{otr}$ | Total traffic induced power consumption of the IoT objects |
The total IoT processing induced power consumption is composed of:

1) The processing induced power consumption of each relay: 
   \[ PC^{pr} = TW_c \cdot RMP \quad \forall c \in R \]  
   \[ PC^{pr} = TW_c \cdot CMP \quad \forall c \in C \]  
   \[ PC^{pr} = TW_c \cdot GMP \quad \forall c \in G \]  
   \[ PC^{pr} = TW_c \cdot UMP \quad \forall c \in ONU \]  
   \[ PC^{pr} = TW_c \cdot LMP \quad \forall c \in OLT \]

The total IoT traffic induced power consumption is composed of:

1) The traffic induced power consumption of each IoT object: 
   \[ PC^{tr} = \sum_{y \in R \cup C \cup G \cup ONU \cup OLT} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot (E^{ot} + \epsilon \cdot d_{2xy}) \quad \forall x \in O \] 

2) The traffic induced power consumption of each relay:
   \[ PC^{tr} = \sum_{y \in R \cup C \cup G \cup ONU \cup OLT} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot (E^{rt} + \epsilon \cdot d_{2xy}) \] 
   \[ + \sum_{y \in ONU \cup OLT} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{rr} \quad \forall x \in R \]

3) The traffic induced power consumption of each coordinator:
   \[ PC^{tr} = \sum_{y \in C} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot (E^{ct} + \epsilon \cdot d_{2xy}) \] 
   \[ + \sum_{y \in R} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{cr} \quad \forall x \in C \]

4) The traffic induced power consumption of each gateway:
   \[ PC^{tr} = \sum_{y \in ONU} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{gr} \] 
   \[ + \sum_{y \in G} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{gtr} \]

5) The traffic induced power consumption of each ONU:
   \[ PC^{tr} = \sum_{y \in ONU} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{u} \] 
   \[ + \sum_{y \in OLT} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{u} \quad \forall x \in ONU \]

6) The traffic induced power consumption of the OLT:
   \[ PC^{tr} = \sum_{y \in ONU} (\lambda_{xy}^{upt} + \lambda_{xy}^{pt}) \cdot E^{i} \] 
   \[ \forall x \in OLT \]

Traffic induced power consumption components of our proposed network are represented by equations (6) to (11). The general structure of these equations is based on radio energy dissipation equation (Friis free-space equation) used in [35]. These equations are comprised of two basic parts the sending and receiving part. Both parts are based on bit rate times the propagation energy per bit. Equation (6) represents the traffic induced power consumption of the IoT objects. This equation considers the sending traffic only because the traffic received by the IoT objects is considered in this model as signalling messages with small data size that can be ignored. On the other hand, equation (11) considers only the receiving traffic induced power consumption of OLT as the OLT layer is the highest layer in the model.
The model objective is to minimize the PON and IoT network power consumption due to traffic processing and aggregation as presented in equation (12). The scaling factor A is introduced to examine the case where the traffic induced power consumption in the networking elements is comparable to their processing induced power consumption.

Objective: Minimize

\[
\sum_{o \in O} P_{\text{tcp}} + \sum_{c \in C} P_{\text{cpp}} + \sum_{r \in R} P_{\text{epr}} + \sum_{v \in V} P_{\text{cvp}} + \sum_{y \in Y} P_{\text{cyp}} + \sum_{l \in L} P_{\text{crl}} + \sum_{o \in O} P_{\text{cop}} + A \left( \sum_{x \in E} P_{\text{ctr}} + \sum_{x \in E} P_{\text{ctu}} \right) + \left( \sum_{x \in E} P_{\text{ctr}} + \sum_{x \in E} P_{\text{ctu}} \right)
\]

(12)

Subject to:

1) IoT network unprocessed traffic constraints

\[
\sum_{v \in VM} \lambda_{oc}^{\text{upt}} = \lambda_{oc}^{\text{upt}}
\]

∀ o ∈ O, ∀ v ∈ VM

(13)

∀ o ∈ O, ∀ c ∈ CN

\[
\sum_{y \in Y} \lambda_{oc}^{\text{cxy}} = \sum_{y \in Y} \lambda_{oc}^{\text{cxy}}
\]

(14)

\[
\lambda_{oc}^{\text{cxy}} = \sum_{c \in CN} \sum_{o \in O} \lambda_{oc}^{\text{cxy}}
\]

∀ x ∈ TN, ∀ y ∈ Ny

(16)

Constraint (13) distributes the unprocessed traffic from IoT objects (o) over a number of VM (v) instances that are hosted in different mini cloudlets (c). It ensures that the total un-processed traffic flows from the IoT object (o) to all VM (v) instances in different mini cloudlets (c) equals to the traffic between that object (o) and the VM (v). Constraint (14) calculates the traffic flowing from IoT objects to each networking element. It ensures that the total un-processed traffic from the IoT object o to all the virtual machines v placed in cloudlet c is equal to the unprocessed traffic from the object o to cloudlet c placed in candidate networking element. Constraint (15) represents the flow conservation for the un-processed traffic from the IoT object o to cloudlet c located in candidate networking element. It ensures that the total un-processed outgoing traffic is equal to the total un-processed incoming traffic for each IoT node except for the source and the destination. Constraint (16) represents the total unprocessed traffic between any IoT node pair (x,y).

2) IoT network processed traffic constraints

\[
\sum_{v \in VM} \lambda_{oc}^{\text{pt}} = F \cdot \sum_{v \in VM} \lambda_{oc}^{\text{upt}}
\]

∀ c ∈ CN

(17)

\[
\sum_{y \in Y} \lambda_{cxy}^{\text{ct}} - \sum_{y \in Y} \lambda_{cxy}^{\text{ct}} = \left\{ \begin{array}{ll}
\lambda_{ct}^{\text{ct}} & \text{if } x = c \\
-\lambda_{ct}^{\text{ct}} & \text{if } x = l \\
0 & \text{otherwise}
\end{array} \right.
\]

∀ c ∈ CN, ∀l ∈ OLT, ∀x ∈ CN: c ≠ l

(18)

\[
\lambda_{xy}^{\text{pt}} = \sum_{c \in CN} \sum_{o \in O} \lambda_{oc}^{\text{pt}}
\]

∀ x ∈ CN, ∀ y ∈ CN

(19)

Constraint (17) calculates the reduced traffic flowing from the candidate networking element host in cloudlet c to the OLT l. Constraint (18) represents the flow conservation for the processed traffic from the candidate networking element hosted cloudlet c to the OLT l. It ensures that the total processed outgoing traffic is equal to the total processed incoming traffic for each IoT and PON node except for the source and the destination. Constraint (19) represents the total processed traffic between any IoT and PON node pair (x,y).

3) Virtual machine placement and workload constraints

\[
\sum_{o \in O} \lambda_{oc}^{\text{upt}} \geq I_{vc}
\]

∀ v ∈ VM, ∀ c ∈ CN

(20)

\[
\sum_{o \in O} \lambda_{oc}^{\text{upt}} \leq \beta \cdot I_{vc}
\]

∀ v ∈ VM, ∀ c ∈ CN

(21)

\[
\sum_{v \in VM} I_{vc} \geq H_{c}
\]

∀ c ∈ CN

(22)

\[
\sum_{v \in VM} I_{vc} \leq \gamma \cdot H_{c}
\]

∀ c ∈ CN

(23)
\[ TW_c = \sum_{v \in VM, c \in CN} W_{bc} \cdot I_{bc} \]  \hspace{1cm} (24)

Constraints (20) and (21) place the virtual machine \( v \) in the cloudlet \( c \) if the cloudlet \( c \) is serving some IoT object requests for this virtual machine. \( \beta \) is a large enough number with units of bps to ensure that \( I_{bc} = 1 \) when \( \sum_{a \in A} I_{bc}^{a} \) is greater than zero, otherwise \( I_{bc} = 0 \). Constraints (22) and (23) build a cloudlet \( c \) in the candidate networking element if this networking element is chosen to host at least one virtual machine \( v \), where \( \gamma \) is a large enough unitless number to ensure that \( H_c = 1 \) if \( \sum_{v \in VM} I_{bc} \) is greater than zero, otherwise \( H_c = 0 \). Constraint (24) calculates the total normalized workload of each built cloudlet \( c \).

III. MILP EVALUATION AND RESULTS

In our evaluation, two IoT networks were considered, supported a PON network. The scenario has in each IoT network: 50 IoT objects, 25 relays, one gateway and a single coordinator. The PON OLT supports two ONUs and each IoT network is connected to an ONU. Figure 2 shows a 30 m × 30 m area which contains the components of each IoT network, namely the IoT objects, relays and coordinator, with 100m separating the gateway from the coordinator. The IoT objects are distributed in the 30 m × 30 m space randomly and uniformly, while the relays are separated by 6 m in a deterministic and uniform fashion as shown in Figure 2. Communication in the IoT network uses the Zigbee protocol which supports the IoT devices. A Gigabit Ethernet link is used to connect the gateway to the ONU. The ONU to OLT fibre link is part of the PON architecture. In our study we only consider the uplink direction as most of the IoT traffic is carried in this direction. As such, we also consider a setup where traffic does not pass from an IoT network to another IoT network through the OLT. We consider in our model the power consumption in the PON modules (ONUs and OLT) due to the traffic flowing in the network. We also consider the power consumption of the IoT network components attributable to transmitters, receivers and power amplifiers which compensate for the propagation losses incorporated in our models [36].

Table III summarizes the parameters used in the model. In terms of power consumption, two parts are considered for each network element in the proposed network; namely the communication and processing parts. The specifications of communication part used in objects, relays and coordinator are based on [37] while we used Cisco 910 industrial router [38] for the communication part of the gateway. In addition we used FTE7502 EPON ONU [39] and FSU7100 EPON OLT [40] as the ONU and OLT elements in the proposed network. The relays, coordinator, gateway, ONU and OLT elements are equipped with Intel Atom Z510 CPU [41] used for processing. We have considered a range of traffic reduction percentages after processing in order to investigate different impacts of processing applications.

![Figure 2 IoT objects and the distribution of relays](image)

| Parameter Name | Value       |
|----------------|-------------|
| Traffic sent from IoT object to a VM type \( \lambda_{LO}^{a} \) | 5 kbps [42] |
| CPU maximum power consumption \( (\text{RMP, CMP, GMP}) \) | 4.64 W[41] |
| Number of CPUs used in a relay, coordinator, gateway, ONU and OLT. | 1, 2, 4, 4, 10 |
| IoT object, relay and coordinator transmitting energy per bit \( (E^{or}, E^{rr}, E^{cf}) \) | 50 nJ/bit [37] |
| Relay and coordinator receiving energy per bit \( (E^{or}, E^{rr}, E^{cr}) \) | 50 nJ/bit [37] |
| Gateway receiving energy per bit \( (E^{gr}) \) | 60 µJ/bit [38] |
| Gateway sending energy per bit \( (E^{gs}) \) | 15 nJ/bit [38] |
| ONU energy per bit \( (E^{s}) \) | 7.5 nJ/bit [40] |
| OLT energy per bit \( (E^{l}) \) | 225.6 pJ/bit [40] |
| Transmission amplifier power coefficient \( (\varepsilon) \) | 255 pJ/(bit.m²) [37] |
| VM type 1 normalized workload in relay, coordinator, gateway, ONU and OLT elements \( (W_{1c}) \) | 0.1, 0.05, 0.025, 0.025, 0.01 [43] |
| VM type 2 normalized workload in relay, coordinator, gateway, ONU and OLT elements \( (W_{2c}) \) | 0.2, 0.1, 0.05, 0.05, 0.02 [43] |
In our evaluation three scenarios were considered. In the first scenario four VM types were considered characterized by heterogeneous VM CPU demands that range from 10% CPU utilization to 40% CPU utilization. Homogeneous, 40%, CPU requirements were considered in the second scenario which has four VM types. In the third scenario, a setup similar to that of Scenario 2 was considered where there are four VM types, all homogeneous, and require 40% of the CPU. The CPU power consumption attributable to traffic at the 10% traffic reduction ratio increases. This is attributed to the smaller traffic volume which induces lower power consumption as the traffic reduction ratio increases. In this case more segments of the network carry the smaller, extracted knowledge, instead of the raw unprocessed traffic.

In Figure 4 also note that Scenario 3 has the lowest power consumption attributable to traffic at the 10% traffic reduction ratio. Scenario 3 is able to place more VMs in the coordinator layer compared to the other scenarios, see Figures 6, 7 and 8 and the 10% traffic reduction case, hence more the knowledge-bearing lower-volume traffic passes to the upper layers. It has to be noted however that this reduction in network power consumption in Scenario 3 is overwhelmed by the increase in CPU power consumption at low reduction percentage, this leading to higher overall power consumption in Scenario 3 compared to the other two scenarios, see Figure 5 at the 10% traffic reduction case. In terms of traffic induced power consumption (at low traffic reduction ratio, i.e., 10%, see Figure 4), the next best is Scenario 1 which places some VMs in the lower network layers (10% case in Figure 6). Scenario 2 places the VMs in the OLT (10% case in Figure 7) which leads to the highest power consumption attributable to traffic (the 10% traffic reduction case in Figure 4).

Examining Figure 3 shows that Scenario 3 has the highest CPU power consumption at the low (10%) traffic reduction percentage. Observe that the OLT has an energy inefficient CPU. This results in the VMs being placed at the lower layers as can be seen in Figure 8 (10% case).

It should be noted that at high traffic reduction ratios (50% - 90%), the VMs are placed in all scenarios in the relay layer in both IoT networks as shown in Figures 6, 7 and 8. This choice results in the minimum power consumption due to traffic as the higher layers are not accessed.

Scenario 1 maintains the lowest CPU power consumption compared to the other two scenarios as it considers heterogeneous VMs. The other two scenarios have comparable CPU power consumption (in the 30% to 90% range in Figure 3) since both scenarios (Scenarios 2 and 3) serve VMs that have similar CPU utilization using the relay elements.

In the second scenario which has four VM types.
relay layer, i.e., the layer closest to the IoT objects, to capitalise on this reduction in traffic. Figure 5 shows that Scenario 1 is the most energy efficient scenario overall. It has the lowest power consumption attributable to processing, and this more than compensates for its higher traffic-induced power consumption. As a result, Scenario 1 has 17% and 19% total power consumption savings compared to Scenarios 2 and 3 respectively.

Figure 3 shows that Scenario 1 is the most energy efficient scenario overall. It has the lowest power consumption attributable to processing, and this more than compensates for its higher traffic-induced power consumption. As a result, Scenario 1 has 17% and 19% total power consumption savings compared to Scenarios 2 and 3 respectively.

![Figure 3 Processing power consumption of the three scenarios](image)

![Figure 4 Traffic power consumption of the three scenarios](image)

![Figure 5 Total power consumption of the three scenarios](image)

IV. EEPIV HEURISTIC

This section validates the MILP model results by presenting the Energy Efficient PON supported IoT Virtualization (EEPIV) heuristic that mimics the MILP model behavior. The pseudo code of the EEPIV heuristic is presented in Figure 9. The heuristic shown in Figure 9 covers all the scenarios of our MILP model as implementing these scenarios relies on changing the input parameters not the constraints that the model is subject to.

The heuristic calculates the total power consumption (TPC) of the network according to the optimum place and the number of mini cloudlets that serve the IoT objects through the hosted VMs. Serving IoT objects by VMs is subject to the limited capabilities of the serving host VM in each cloudlet as below:
There should be sufficient processor capacity in each candidate cloudlet to accommodate the hosted VM workload. The intended VM $v$ that is requested by IoT object $o$ in each network should not have been hosted by any other cloudlet in this network before.

If all the serving constraints above are met, then the heuristic hosts the intended VM in the candidate cloudlet to satisfy the IoT object request and sets the binary indicator $F_{cv}$ accordingly. The total workload of each hosted cloudlet in the candidate place is calculated depending on the binary indicator $F_{cv}$. Since the processing induced power consumption of each processing element is a function of the total workload of the cloudlet, the heuristic calculates the processing induced power consumption of all the processing elements in the proposed network (relays, coordinators, gateways, ONUs and OLT) as shown in steps 11 to 35 in Figure 9. The end-to-end traffic generated by the IoT objects’ requests is next calculated by the heuristic. The traffic passes through two stages: the first stage flows from the generator (IoT object) to the destined VM which is represented by $\lambda_{oc}$ (unprocessed traffic). The second stage comes after the processing stage. In this stage, the processed traffic $\lambda_{ct}$ (reduced traffic) flows from the cloudlet to the last layer in the network which is represented in our proposed network by the OLT layer. The intermediate traffic between each node pair in the network is calculated by the heuristic model based on the end to end traffic. The heuristic routes the traffic through these intermediate nodes from the source to the destination using a minimum hop algorithm to reduce the traffic induced power consumption. Finally, the heuristic calculates the total power consumption $TPC$ by summing all the processing and traffic induced power consumption of all nodes.

Inputs: $VM = \{1 ... NVM\}$
$CN = \{1 ... NCN\}$
$O = \{1 ... NO\}$
$R = \{1 ... NR\}$
$C = \{1 ... NC\}$
$G = \{1 ... NG\}$
$ONU = \{1 ... NONU\}$
$OLT = \{1 ... NOLT\}$

Output: No. of Served Objects
Total Power Consumption (TPC)

1. For each candidate cloudlet that can host a required VM $c \in CN$ Do

2. For each Virtual Machine required by an object $v \in VM$ Do

3. If $U_{cv} > 0$ Then

4. If all serving constraints are met Then

5. $F_{cv}(c, v) = 1$

6. Calculate the workload of the hosting cloudlet $TCW_c$

7. End If

8. End If

9. End For

10. End For

11. For Each relay $(r \in R)$ Do

12. If the hosting cloudlet is placed in relay layer $R \ c \in CN$ Do

13. Calculate $R_{PPC}$

14. End If

15. End For

16. For Each coordinator $(c \in C)$ Do

17. If the hosting cloudlet is placed in coordinator layer $C \ c \in CN$ Do

18. Calculate $C_{PPC}$

19. End If

20. End For

21. For each gateway $(g \in G)$ Do

22. If the hosting cloudlet is placed in gateway layer $c \in CN$ Do

23. Calculate $G_{PPC}$

24. End If

25. End For

26. For each ONU $(u \in ONU)$

27. If the hosting cloudlet is placed in ONU layer $c \in CN$ Do

28. Calculate $PC_{up}$

29. End If

30. End For

31. For each OLT $(l \in OLT)$

32. If the hosting cloudlet is placed in OLT layer $c \in CN$ Do

33. Calculate $PC_{up}$
34. **End If**
35. **End For**
36. **For** each IoT object served by a cloudlet \( o \in O \) **Do**
37. **For** each hosting cloudlet \( c \in CN \) **Do**
38. Calculate end to end traffic that flows from each object to the cloudlet that serves this object
39. **End For**
40. **End For**
41. **For** each hosting cloudlet \( c \in CN \) **Do**
42. **For** each OLT \( (l \in OLT) \) **Do**
43. Calculate the end to end reduced traffic from the cloudlet to the OLT
44. **End For**
45. **End For**
46. **For** each IoT object \( o \in O \) **Do**
47. Calculate \( TO_{-tr} \)
48. **End For**
49. **For** each relay \( (r \in R) \) **Do**
50. Calculate \( TR_{-tr} \) based on minimum hop path between node pair \( (x,y) \)
51. **End For**
52. **For** each coordinator \( (c \in C) \) **Do**
53. Calculate \( TC_{-tr} \) based on minimum hop path between node pair \( (x,y) \)
54. **End For**
55. **For** each gateway \( (g \in G) \) **Do**
56. Calculate \( TG_{-tr} \) based on minimum hop path between node pair \( (x,y) \)
57. **End For**
58. **For** each ONU \( (u \in ONU) \) **Do**
59. Calculate \( ONU_{-tr} \) based on minimum hop path between node pair \( (x,y) \)
60. **End For**
61. **For** each OLT \( (l \in OLT) \) **Do**
62. Calculate \( ONU_{-tr} \) based on minimum hop path between node pair \( (x,y) \)
63. **End For**
64. Calculate total power consumption
\[
TPC = \sum_{r \in R} R_{PPC} + \sum_{c \in C} C_{PPC} + \sum_{g \in G} G_{PPC} + \sum_{u \in ONU} ONU_{PPC} + \sum_{o \in O} TO_{-tr} + \sum_{r \in R} TR_{-tr} + \sum_{c \in C} TC_{-tr} + \sum_{g \in G} TG_{-tr} + \sum_{u \in ONU} ONU_{-tr} + \sum_{l \in OLT} OLT_{-tr}
\]

**Figure 9** pseudo code of EEPIV heuristic

**V. EEPIV HEURISTIC RESULTS**

We used the same inputs in Table III for the heuristic. The heuristic results show close agreement with the MILP results comparing Figure 10 with Figure 5.

**Figure 10** Total power consumption of the three scenarios in the heuristic

Scenarios 1 and 2 result in lower processing induced power consumption in MILP than in the heuristic at low reduction percentages (10%, Figures 3 and 11). This results from placing/using more VM copies in the heuristic (8 VMs) than in the MILP as shown in Figures 6 and 7.
Scenarios 1 and 2 in heuristic result in lower traffic induced power consumption than in MILP at traffic reduction percentages of 10% (Figures 4 and 12). This results from the MILP placing the serving VMs in cloudlets at higher layers (Figures 6 and 7) while all the cloudlets in the heuristic are distributed throughout the lower layer (relay layer). Placing cloudlets in higher layers results in sending more unprocessed traffic (unreduced traffic) to higher layers which in turn results in higher traffic induced power consumption. However, all the scenarios in the heuristic consume higher traffic induced power than in the MILP for the rest of the reduction percentage values as a result of the different distribution of the cloudlets in the proposed network. Since for each cloudlet, the heuristic attempts to place it in the first network element that can accommodate this cloudlet, the heuristic placed all the cloudlets in the relay layer without consideration of the closeness of the cloudlet to the IoT objects. On other side, the MILP places the cloudlets in an optimum way to minimize the traffic and processing induced power consumed by all elements of the proposed network.

The number of cloudlets placed using the heuristic (2 cloudlets in Scenario 1, and 4 cloudlets in Scenarios 2 and 3 for all traffic reduction percentage values with 8 VMs in all scenarios) is less than in the MILP as one of main processes in the heuristic is bin packing where VMs must be packed into a finite number of bins (cloudlets) in a way that minimizes the number of bins used and hence the processing power consumption.

VI. CONCLUSIONS
We have developed a MILP model to optimize the placement of processing tasks in IoT networks supported by a PON infrastructure in order to minimize the overall power consumption. The model optimizes the number and location of cloudlets that are created to host the VMs where the tasks are processed. The total power consumption is made up of the power consumption due to processing and the power consumption due to traffic routing. The IoT data is reduced in volume after processing which leads to lower traffic induced power consumption. This traffic reduction after processing further influences the placement of VMs at different traffic reduction ratios. If the traffic reduction ratio is small, our results indicate that the best location to place the processing, ie the VMs is a location that can be shared by all or the majority of the IoT devices. In our architecture, this location is the OLT, which in the case consolidates the processing offered by the VMs. At the other end, when the traffic reduction ratio is high, ie if the traffic after processing is much smaller than the traffic before processing, our results show that the VMs (ie the processing) has to be placed as close to the IoT objects as possible. In our architecture this is the relay layer. Our results indicate that a power saving of up to 19% can be obtained by placing the processing (VMs) at the appropriate locations in the given setup. We have developed a heuristic for the placement of VMs and routing of information. It serves two purposes. Firstly, the heuristic is simple and therefore enables fast operation. Secondly, it acts to verify the MILP where we observed very close agreement between the heuristic and the MILP. In particular, Scenario 1 has resulted in power savings when using the heuristic of 17% (MILP 17%) and 17% (MILP 19%) compared to Scenarios 2 and 3.

ACKNOWLEDGMENTS
We would like to acknowledge funding from the Engineering and Physical Sciences Research Council (EPSRC) for the INTERNET (EP/H040536/1), STAR (EP/K016873/1) and TOWS (EP/S016570/1) projects. The first author Dr. Zaineb Al-Azez would like to thank Dr. Ahmed Al-Quzweeni for many helpful discussions and the Higher Committee for Education Development in Iraq (HCED) for funding her PhD scholarship. All data are provided in full in the results section of this paper.

REFERENCES
[1] A. Beloglazov, R. Buyya, "Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers", Proc. 8th Int'l Workshop MGC, pp. 1-6, 2010, [online] Available: http://portal.acm.org/citation.cfm?doid=1890799.1890803 [Last accessed 26 Jun. 2018]. J. Pan , and J. McElhannon, "Future Edge Cloud and Edge Computing for Internet of Things Applications", in IEEE Internet of Things Journal, vol. 5, no. 1, pp. 439-449, Feb. 2018.
[2] Zhou, Wei, Yan Jia, Anni Peng, Yuqing Zhang, and Peng Liu, "The Effect of IoT New Features on Security and Privacy: New Threats, Existing Solutions, and Challenges Yet to Be Solved," IEEE Internet of Things Journal (early access), 2018.
[3] J. Pan, and J. McElhannon, "Future Edge Cloud and Edge Computing for Internet of Things Applications", in *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 23-38, 2018.

[4] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," in *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347-2376, Fourthquarter 2015.

[5] S. H. Shah and I. Yaqoob, "A survey: Internet of Things (IOT) technologies, applications and challenges," 2016 IEEE Smart Energy Grid Engineering (SEGE), Oshkosh, ON, 2016, pp. 381-385.

[6] R. Khan, S. U. Khan, R. Zaher and S. Khan, "Future internet: the internet of things architecture, possible applications and key challenges," 10th International Conference on Frontiers of Information Technology, pp. 257-260, 2012.

[7] A.Q. Lawey, T. E. H. El-Gorashi and J. M. H. Elmirghani, "Energy efficient cloud networks over core networks," *IEEE/OSA J. Lightwave Technology*, vol. 32, No. 7, pp. 1261-1281, 2014.

[8] H.M.M., Ali, A.Q. Lawey, T.E.H. El-Gorashi, and J.M.H. Elmirghani, "Future Energy Efficient Data Centers With Disaggregated Servers," *IEEE/OSA Journal of Lightwave Technology*, vol. 35, No. 24, pp. 5361 – 5380, 2017.

[9] X. Dong, T. El-Gorashi, and J. Elmirghani, "Green IP Over WDM Networks With Data Centers," Lightwave Technology, Journal of, vol. 29, no. 12, pp. 1861-1880, June 2011.

[10] N. I. Osman, T. El-Gorashi, E. Krue, and J. M. H. Elmirghani, "Energy Efficient Future High-Definition TV,” *Journal of Lightwave Technology*, vol. 32, no. 13, pp. 2364-2381, July 2014.

[11] A.Q. Lawey, T.E.H. El-Gorashi, and J.M.H. Elmirghani, "BitTorrent Content Distribution in Optical Networks,” *IEEE/OSA Journal of Lightwave Technology*, vol. 32, No. 21, pp. 3687 – 3693, 2014.

[12] L. Nonde, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Energy efficient virtual network embedding for cloud networks," *IEEE/OSA j. Lightwave technology*, Vol. 33, No. 9, pp. 1828-1849, 2015.

[13] A.N. Al-quzaweeni, A.Q. Lawey, T.E.H. El-Gorashi, and J.M.H. Elmirghani, "Optimized Energy Aware 5G Network Function Virtualization,” *IEEE Access*, vol. 7, pp. 44939 - 44958, 2019.

[14] Al-Azez, Z., Lawey, A., El-Gorashi, T.E.H., and Elmirghani, J.M.H., “Energy Efficient IoT Virtualization Framework with Peer to Peer Networking and Processing” *IEEE Access*, vol. 7, pp. 50697 - 50709, 2019.

[15] X. Dong, T. E. H. El-Gorashi and J. M. H. Elmirghani, “On the energy efficiency of physical topology design for IP over WDM networks,” *IEEE/OSA j. Lightwave technology*, Vol. 30, No. 12, pp. 1931-1942, 2012.

[16] J. M. H. Elmirghani, T. Klein, K. Hinton, L. Nonde, A. Q. Lawey, T. E. H. El-Gorashi, M. O. I. Musa, and X. Dong, “Greentouch GreenMeter core network energy-efficiency improvement measures and optimization,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A250-A269, Feb 2018.

[17] M. Musa, T.E.H. El-Gorashi and J.M.H. Elmirghani, “Bounds on Greentouch GreenMeter Network Energy Efficiency,” *IEEE/OSA Journal of Lightwave Technology*, vol. 36, No. 23, pp. 5395-5405, 2018.

[18] B. Bathula, M. Alresheedi, and J.M.H. Elmirghani, “Energy efficient architectures for optical networks,” Proc IEEE London Communications Symposium, London, Sept. 2009.

[19] B. Bathula, and J.M.H. Elmirghani, “Energy Efficient Optical Burst Switched (OBS) Networks,” *IEEE GLOBECOM ’09*, Honolulu, Hawaii, USA, November 30-December 04, 2009.

[20] X. Dong, A.Q. Lawey, T.E.H. El-Gorashi, and J.M.H. Elmirghani, “Energy Efficient Core Networks,” Proc 16th *IEEE Conference on Optical Network Design and Modelling (ONDM’12),* 17-20 April, 2012, UK.

[21] X. Dong, T.E.H. El-Gorashi and J.M.H. Elmirghani, “Green optical orthogonal frequency-division multiplexing networks,” *IET Optoelectronics*, vol. 8, No. 3, pp. 137 – 148, 2014.

[22] X. Dong, T. El-Gorashi, and J. Elmirghani, “IP Over WDM Networks Employing Renewable Energy Sources,” *Lightwave Technology, Journal of*, vol. 29, no. 1, pp. 3-14, Jan 2011.

[23] M. Musa, T.E.H. El-Gorashi and J.M.H. Elmirghani, “Bounds for Energy-Efficient Survivable IP Over WDM Networks with Network Coding,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 5, pp. 471-481, 2018.

[24] M. Musa, T.E.H. El-Gorashi and J.M.H. Elmirghani, “Energy Efficient Survivable IP-Over-WDM Networks With Network Coding,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, No. 3, pp. 207-217, 2017.
Zaineb T. Al-Azez received the B.Sc. and M.Sc. degrees in computer engineering from Nahrian University, Baghdad, Iraq, in 2005 and 2008, respectively, and the Ph.D. degree in communication networks from the School of Electronic and Electrical Engineering, University of Leeds, U.K., in 2018. From 2009 to 2014, she was an Assistant Lecturer with Al-Mansour University College, Baghdad. Her current research interests include energy efficiency in the Internet of Things networks, cloud computing, and virtualization.

Ahmed Q. Lawey received the B.Sc. and M.Sc. degrees (Hons.) in computer engineering from Nahrian University, Iraq, in 2002 and 2005, respectively, and the Ph.D. degree in communication networks from the University of Leeds, U.K., in 2015. From 2005 to 2010, he was a Core Network Engineer with ZTE Corporation for Telecomunication, Iraq Branch. He is currently a Lecturer in communication networks with the School of Electronic and Electrical Engineering, University of Leeds. His current research interests include energy efficiency in optical and wireless networks, big data, cloud computing, and the Internet of Things.

Taisir EL-Gorashi received the B.S. degree (first-class Hons.) in electrical and electronic engineering from the University of Khartoum, Khartoum, Sudan, in 2004, the M.Sc. degree (with distinction) in photonic and communication systems from the University of Wales, Swansea, UK, in 2005, and the PhD degree in optical networking from the University of Leeds, Leeds, UK, in 2010. She is currently a Lecturer in optical networks in the School of Electrical and Electronic Engineering, University of Leeds. Previously, she held a Postdoctoral Research post at the University of Leeds (2010–2014), where she focused on the energy efficiency of optical networks investigating the use of renewable energy in core networks, green IP over WDM networks with datacenters, energy efficient physical topology design, energy efficiency of content distribution networks, distributed cloud computing, network virtualization and Big Data. In 2012, she was a BT Research Fellow, where she developed energy efficient hybrid wireless-optical broadband access networks and explored the dynamics of TV viewing behavior and program popularity. The energy efficiency techniques developed during her postdoctoral research contributed 3 out of the 8 carefully chosen core network energy efficiency improvement measures recommended by the GreenTouch consortium for every operator network worldwide. Her work led to several invited talks at GreenTouch, Bell Labs, Optical Network Design and Modelling conference, Optical Fiber Communications conference, International Conference on Computer Communications, EU Future Internet Assembly, IEEE Sustainable ICT Summit and IEEE 5G World Forum and collaboration with Nokia and Huawei.

Jaafar M. H. Elmirghani is the Director of the Institute of Communication and Power Networks within the School of Electronic and Electrical Engineering, University of Leeds, UK. He joined Leeds in 2007 and prior to that (2000–2007) as chair in optical communications at the University of Wales Swansea he founded, developed and directed the Institute of Advanced Telecommunications and the Technium Digital (TD), a technology incubator/spin-off hub. He has provided outstanding leadership in a number of large research projects at the IAT and TD. He received the Ph.D. in the synchronization of optical systems and optical receiver design from the University of Huddersfield UK in 1994 and the DSc in Communication Systems and Networks from University of Leeds, UK, in 2012. He has co-authored Photonic Switching Technology: Systems and Networks, (Wiley) and has published over 500 papers. He has research interests in optical systems and networks. Prof. Elmirghani is Fellow of the IET, Fellow of the Institute of Physics and Senior Member of IEEE. He was Chairman of IEEE Comsoc Transmission Access and Optical Systems technical committee and was Chairman of IEEE Comsoc Signal Processing and Communications Electronics technical committee, and an editor of IEEE Communications Magazine. He was founding Chair of the Advanced Signal Processing for Communication Symposium which started at IEEE GLOBECOM’99 and has continued since at every ICC and GLOBECOM. Prof. Elmirghani was also founding Chair of the first IEEE ICC/GLOBECOM optical symposium at GLOBECOM’00, the Future Photonic Network Technologies, Architectures and Protocols Symposium. He chaired this Symposium, which continues to date under different names. He was the founding chair of the first Green Track at ICC/GLOBECOM at GLOBECOM 2011, and is Chair of the IEEE Sustainable ICT Initiative within the IEEE Technical Activities Board (TAB) Future Directions Committee (FDC) and within the IEEE Communications Society, a pan IEEE Societies Initiative responsible for Green and Sustainable ICT activities across IEEE, 2012-present. He is and has been on the technical program committee of 39 IEEE ICC/GLOBECOM conferences between 1995 and 2020 including 19 times as Symposium Chair. He received the IEEE Communications Society Hal Sobol award, the IEEE Comsoc Chapter Achievement award for excellence in chapter activities (both in 2005), the University of Wales Swansea Outstanding Research Achievement Award, 2006, the IEEE Communications Society Signal Processing and Communication Electronics outstanding service award, 2009, a best paper award at IEEE ICC’2013, the IEEE Comsoc Transmission Access and Optical Systems outstanding Service award 2015 in recognition of “Leadership and Contributions to the Area of Green Communications”.
received the GreenTouch 1000x award in 2015 for “pioneering research contributions to the field of energy efficiency in telecommunications”, the 2016 IET Optoelectronics Premium Award and shared with 6 GreenTouch innovators the 2016 Edison Award in the “Collective Disruption” Category for their work on the GreenMeter, an international competition, clear evidence of his seminal contributions to Green Communications which have a lasting impact on the environment (green) and society. He is currently an editor of: IET Optoelectronics, Journal of Optical Communications and was editor of IEEE Communications Surveys and Tutorials and IEEE Journal on Selected Areas in Communications series on Green Communications and Networking. He was Co-Chair of the GreenTouch Wired, Core and Access Networks Working Group, an adviser to the Commonwealth Scholarship Commission, member of the Royal Society International Joint Projects Panel and member of the Engineering and Physical Sciences Research Council (EPSRC) College. He was Principal Investigator (PI) of the £6m EPSRC INtelligent Energy awaRe NETworks (INTERNET) Programme Grant, 2010-2016 and is currently PI of the £6.6m EPSRC Terabit Bidirectional Multi-user Optical Wireless System (TOWS) for 6G LiFi Programme Grant, 2019-2024. He has been awarded in excess of £30 million in grants to date from EPSRC, the EU and industry and has held prestigious fellowships funded by the Royal Society and by BT. He was an IEEE Comsoc Distinguished Lecturer 2013-2016.