Abstract—Integrated optical and wireless networks are considered as the next generation access networks because of the potential benefits from both technologies. In this paper, we consider the integrated BS/ONU (IBO) placement problem in the hybrid EPON-WiMAX system to efficiently utilize network resources as well as reduce the system cost. We propose a modified clustering algorithm (MCA) to obtain the near-optimal result that minimizes the number of IBOs needed to support all wireless relay BSs residing in the wireless part of the integrated system, while satisfying several necessary quality-of-service (QoS) constraints, e.g., hop count, cluster size, etc. In contrast to the existing related work, our MCA approach forms the clusters starting from the network edge towards its center and the construction of clusters is not only based on the “greedy” idea but also considers the load balance. Simulation and numerical results show that the number of IBOs generated from our algorithm is always less than or equal to the best results known in the current literature, but our algorithm can achieve better network performance in terms of the reduced average transmission delay, lowered link failure probability and more balanced load.

Index Terms—EPON, WiMAX, Integrated system, placement problem, clustering algorithm, QoS constraints

I. INTRODUCTION

In order to meet the users’ increasing demand for higher bandwidth requirement, provide broadband and ubiquitous high-speed internet access, and address the growing gap between core network and local area network (“last mile” problem) effectively, research on the hybrid optical and wireless networks has been studied in the recent years [1], [2], [3], [4], [5], [16]. The EPON-WiMAX integrated system has attracted a lot of interests from both academia and industry [4], [5]. EPON is an Ethernet-based passive optical network that only uses passive equipments, e.g. splitter, to reduce the network maintenance cost. As a promising wireless solution to broadband access, EPON provides much more reliable transmission and much higher bandwidth with several Gbps compared with wireless network as well as covering long distance (around 20 kilometers). However, the fixed infrastructure of EPON not only limits its coverage in a high densely populated area due to the high fiber layout cost, but also makes it hard to deploy in certain rugged environment. WiMAX, another promising wireless solution to broadband access, supports mobility and ubiquitous access for end users in the metropolitan area, but suffers vulnerable wireless transmission and provides limited bandwidth around several hundred of Mbps. Thus, integrating the two different technologies together can utilize their complementary advantages to provide better service for end users and increase revenue for the service providers.

Fig. 1 illustrates the hybrid EPON-WiMAX system architecture. EPON uses a tree topology where the optical line terminal (OLT) located at the telecom central office (CO) connects multiple optical network units (ONUs) through the passive splitter. In the wireless part, WiMAX base stations (BSs) are grouped into clusters. Within each cluster, one WiMAX BS is selected to combine together with one ONU (BSs) are grouped into clusters. Within each cluster, one WiMAX BS is selected to combine together with one ONU forming the integrated point of the hybrid network and the rest of WiMAX BSs in the cluster referred as relay stations form a multi-hop wireless mesh network. These integrated points called integrated BS/ONUs (IBOs) are the locations where the wireless part and optical part meet together. For data transmission, an end user sends packets to its closest relay station and this relay station forwards the packets to the corresponding IBO through one hop or multi-hops in the cluster. The IBO will forward the packets to the OLT through the optical connection and then to the core network. Then the core network will direct the packets to the receiver through the proper access network, such as 3G cellular, WLAN, WiMAX, etc. If the receiver resides in the hybrid EPON-WiMAX network, the data flows in the reverse direction of transmission.

When developing the hybrid EPON-WiMAX network, we need to address several issues including placement of IBOs, resource allocation and scheduling, routing protocol design, etc., in order to make the whole system working efficiently with the minimum system cost. In this paper, we mainly focus on the IBO placement problem. Given a WiMAX network topology, we want to minimize the number of IBOs to lower the fiber layout cost, equipments cost and installation cost, while still maintaining the network connectivity and satisfying the quality-of-service (QoS) requirement to meet the user demand. We propose a modified clustering algorithm (MCA) to solve this problem as well as take the required QoS constraints into account. Our MCA approach forms clusters starting from the network edge towards the center, constructs
each cluster based on the QoS constraints as well as a predefined condition, and then performs shift operation as well as updating the IBO final location. In general, our approach can minimize the number of clusters, reduce average transmission delay, and balance load. Numerical and simulation results verify that our proposed algorithm is better than others presented in the literature in most case, or at least the same in some cases.

The rest of paper is organized as follows: Section II reviews more related work. Section III describes the system model and assumptions. In section IV, a linear programming optimization problem and the proposed MCA approach are discussed in detail. Section V presents some simulation and numerical results. Section VI gives the conclusion remarks.

II. RELATED WORK

OUN placement in hybrid optical-wireless broadband access networks has been studied in [5], [9], [10], [16], using greedy algorithm, simulated annealing algorithm, and combined heuristic algorithm. The former two algorithms, given the number of ONUs and user locations, they aim to find out the optimal ONU locations through minimizing some cost functions, which are usually formulated as the average distance between end users and ONUs to represent the average transmission delay and fiber layout cost. The third algorithm is similar to the former ones, but obtaining the number of Base Stations first based on the co-channel interference threshold, then find out the optimal solution based on the greedy algorithm. However, they do not address how to set the number of ONUs. Thus, at the very beginning stage, the desired ONU placement should use minimum number of ONUs to obtain the minimum cost and then optimize the locations of ONUs to meet the QoS requirements.

Note that the functions of IBO are similar to those of sink nodes in wireless sensor network and gateways in the wireless mesh network. Thus, the work on the placement of sink nodes or gateways is also related. [6], [7] and [8] have studied sink node deployment in wireless sensor networks. Given the number of sink nodes and sensor locations, they use popular algorithms such as integer linear programming (ILP), genetic algorithm or k-mean clustering algorithm to find out the optimal locations of the sink nodes. Similar to [9], [10], they aim to shorten the average Euclidean distance between sensor node and sink node to save the energy of the sensor nodes consumed when relaying data packets in such multi-hop wireless sensor network.

For the gateway placement in wireless mesh networks, literatures in [11], [12], [13], [14], [15] have studied how to minimize the number of gateways taking into account several QoS constraints, e.g., hop distance, cluster size, etc, given the network topology. The earliest work [11] breaks the optimization problem into two sub-problems and uses Dominating Independent Set approach to solve it. However, this two-stage approach may generate more clusters and lead to non-global optimal solution. Also, the work assumes the network MAC (medium access control) layer protocol using TDMA (time-division-multiplexing-access) scheme which requires synchronization, is hard to achieve in large scale mesh networks, so limits its application. In [12], the authors formulate three different link models and propose a greedy algorithm forming clusters iteratively to maximize the traffic demand, but the network delay performance will be the trade-off. In [13], the authors choose the cluster-head and form the cluster in parallel, which will have less number of gateways than the result obtained from [11]. When the QoS constrains are violated, the algorithm breaks the big cluster into two small ones in order to satisfy the QoS requirement, but this will result in more clusters. In [14], an IGW-rooted tree approach is used to select the internet gateway (IGW) and form the cluster. But it only deals with one IGW selection case and does not specify after forming one cluster, which node should be the first to select the next IGW. Also, cluster-based algorithm in [11-14] will not achieve the global optimum, as we will see later.

The most similar work to ours is the work of Drabu [15]. In [15], the author proposes a split-merge-shift (SMS) algorithm to minimize the number of clusters. This algorithm forms one-hop cluster first at the node with the maximum node degree. Then it merges neighboring clusters with small cluster size. When merging operation cannot works, it splits small cluster into singleton clusters and uses shift operation to merge singleton clusters into neighboring large clusters to minimize the number of gateways. Till now, [15] has been proved to be the best work to get the minimum number of gateways in wireless mesh network. However, in the shift operation in [15], the authors do not specify when multiple exchange paths are available, which path should be selected and which node should be shifted. These ambiguities will result in different formation of clusters. Also, when performing the merge and shift operation, the algorithm does not specify whether to relocate the gateway position and this will affect the final result of the number of clusters. At last, after the algorithm stops, the authors do not consider updating the gateway location to optimize average transmission delay and fault-tolerance.

III. SYSTEM MODEL AND ASSUMPTIONS

A. System Model

Our system model is shown in Fig. 2. We consider the integrated IBO placement problem for hybrid EPON-WiMAX system where WiMAX forms wireless mesh network (WMN). We model an n - node WMN as an undirected connected graph \(G(V,E)\), where \(V\) is the vertex set represents...
the set of nodes including WiMAX relay stations and
integrated points IBOs and $E$ is the set of all the edges
representing the communication link between two neighboring
nodes. Two nodes are called neighbors if and only if their
Euclidean distance $d(u,v)$ is less than or equal to the node
transmission range $r_t$. The neighbors of a node $v$, denoted by
$v_N$, is the set of nodes located within node $v$’s transmission
range. The number of neighbors of a node $v$ is called the
degree of $v$, denoted by $\delta$. The hop distance between any two
nodes $u$ and $v$, denoted by $(u,v)$, is the minimum number
of hops between them.

We assume the system has uniformly distributed traffic on
each node and there are no peer-to-peer transmissions between
relay BSs. Each node has the same transmission range.

For convenience, we use the adjacent matrix to represent
the $n$-node graph $(V,G)$, where,

$$A[i,j]=\begin{cases} 1 & \text{if } d(v_i,v_j) \leq r_t, \quad 1 \leq i,j \leq n; \ v_i,v_j \in V. \\ 0 & \text{if } d(v_i,v_j) > r_t. \end{cases} \quad (1)$$

### B. QoS constraints

In our system, we consider two kinds of QoS constraints,
including hop distance and cluster size and specify them as follows:

- Maximum Hop count ($R_h$): We know that the less the
  hop count is, the smaller the transmission delay and link
  failure probability are. Thus, setting an upper bound on the
  hop count is equivalent to allowing certain maximum transmission delay and certain
  maximum link failure probability. The hop distance constraint is specified as:

$$h(g_i,v) \leq R_h, \quad \forall v \in G_i \quad \text{and} \quad v \neq g_i \quad (2)$$

where $G_i$ is cluster $i$, $g_i$ is the IBO node in $G_i$, and
$h(g_i,v)$ is the minimum hop distance between node $v$
and IBO $g_i$.

- Maximum cluster size ($C_{\text{max}}$): to guarantee that in
each cluster, the total traffic demand should not exceed
the IBO maximum capacity, and this is formulated as:

$$|G_i| \leq C_{\text{max}} \quad \forall G_i \in \psi \quad (3)$$

where $|G_i|$ is the number of nodes inside a cluster and
$\psi$ is the union of all the clusters that covers the whole
system.

### IV. IBO PLACEMENT ALGORITHMS

#### A. Integer Linear Program Formulation

Let $N = V$ be set of WiMAX relay stations and $G \subseteq N$ be
the set of IBOs which are the subset of $N$. We introduce a
binary integer $y_i$ to indicate whether node $i$ has been selected
as IBO. In order to represent the relationship between a relay
station and an IBO, we define another binary variable $x_{i,j}$. If
this variable equal to 1, this means that the relay station
$j \in N$ is assigned to IBO $i$; otherwise, $x_{i,j} \text{ equal to 0. Thus, we can}
formulate the IBO placement problem subject to the
considered QoS constraints as an optimization problem as
follows:

$$\min \sum_{i \in N} y_i$$

Subject to:

(a) \ \forall j \in N: \ \sum_{i \in N} x_{i,j} = 1

(b) \ \forall i, j \in N: y_i \geq x_{i,j}

(c) \ \forall i, j \in N: h_{i,j} \cdot x_{i,j} \leq R_h

(d) \ \forall i \in N: \ \sum_{j \in N} x_{i,j} \leq C_{\text{max}}

(e) \ \forall i, t \in N: x(i,j) + x(j,i) \leq 1, \ i \neq j

(f) \ \forall i, j, k \in N: x(i,j) \leq \sum_{k \in N} x(i,k)

(g) \ \forall i \in N: y_i \in \{0,1\}

(h) \ \forall i, j \in N: x_{i,j} \in \{0,1\}

Condition (a) denotes that each relay station is assigned to
one and only one IBO, so this guarantees the whole network
coverage and generates non-overlap clusters. Inequality (b)
means that a relay station should be assigned to an IBO after
that IBO has been set up. Inequalities (c), and (d) satisfy the
QoS constraints in terms of the hop count and cluster size.
Inequality (e) ensures that if a relay station has been assigned
to an IBO, then this relay station can not be used as a
candidate of IBO. Inequality (f) ensures that a relay station can
only be assigned to an IBO if at least one of the parents of this
relay station in the shortest path tree rooted at this IBO has
already been assigned to this IBO. Inequalities (g) and (h)
means that the variables here only can take binary value. The
optimization problem can be solved via Integer Linear
Programming (ILP), which, however, is NP-hard, proved in
[12], [13]. Thus, we need to provide a heuristic algorithm to
obtain the near-optimal result, which is described in detail next.

B. MCA approach

The SMS algorithm in [15] forms clusters from the node with the maximum node degree and then performs the split-merge-shift operations without relocating the gateway location. This generally will result in the clusters with unbalanced load or high hop distance, that is, some clusters may have very dense nodes with one-hop distance to the gateway around the network center and some clusters may have just a few nodes but with high hop distance around the network edge. In contrast to the SMS algorithm, in our MCA approach, we form the clusters from the network edge towards its center and update the IBO location at the final stage to better balance the load and hop count among clusters. The proposed MCA approach is composed of the following four components:

- Select a node as the starting point to form a cluster, given in Procedure 1.
- Cluster forming criterion, given in Procedure 2.
- Shift operation used to reduce the number of clusters after cover all the nodes at the first time, given in Procedure 3.
- Update the IBO at the final stage, given in Algorithm 1.

The detailed MCA description is illustrated in Algorithm 1.

![Algorithm 1](https://example.com/algorithm1.png)

![Procedure 1 – starting node selection](https://example.com/proc1.png)

![Procedure 2 – cluster construction](https://example.com/proc2.png)

![Procedure 3 – shift operation](https://example.com/proc3.png)

C. Algorithms illustration

We use a 40-node network to illustrate the MCA approach implementation and the initial network topology randomly generated is shown in Fig. 3(a). In this example we set $R_g = 3$ and $C_{max} = 8$.

Node 1 is selected as the starting point to form the first cluster according to Step a, Procedure 1. Then node 2 is selected as the initial IBO since it has the maximum node degree $\delta = 6$ within node 1’s 2-hop range. When forming the cluster, node 1 is first added into the cluster according to Step a, Procedure 2.

Then node 2’s 1-hop neighbor node 7 is added in since it has the smallest node degree $\delta' = 1$. Node 6, is added since it has the maximum node degree $\delta = 6$ within node 1’s 2-hop range. When forming the cluster, node 1 is first added into the cluster according to Step a, Procedure 2.

![Node 1 selected as the starting point](https://example.com/node1.png)

![Node 2 selected as the initial IBO](https://example.com/node2.png)

![Node 7 added in since it has the smallest node degree](https://example.com/node7.png)

![Node 6 added since it has the maximum node degree](https://example.com/node6.png)

The detailed MCA description is illustrated in Algorithm 1.
form the next cluster, and the algorithm continues to form clusters based on Algorithm 1 until the stop condition in Algorithm 1, step 3 is satisfied. This result is shown in Fig. 3(b).

Now we move to Step 4 in Algorithm 1, and perform the shift operation to reduce the number of clusters. Based on $\bar{C}_2 = \frac{1}{M_2} \sum_{i=1}^{M_2} C_i$ from Step a, Procedure3, we find that each cluster should have the same number of nodes which equal to 8. Thus, we need to shift one node from cluster 4 or 5 into cluster 1 first. According to step d, Procedure 3, we shift two-hop node 8 from cluster 4 to cluster 1 rather than shift one-hop nodes 9 and 18 to reduce the average transmission delay. Following the same criterion, node 16 is shifted from cluster 5 into cluster 4 and nodes 22, 23 and 24 in cluster 6 are shifted into cluster 5. Then we get the minimum number of clusters 5 to cover all the nodes, and this is shown in Fig. 3(c).

Finally, we update the IBO location in each cluster according to $\forall k \in G_i : H(k) = \arg\min \left\{ \sum_{j \in G_i} h(k,j), \ k \neq j \right\}$ in Step 5, Algorithm 1, and this final result is shown in Fig. 3(d).

V. SIMULATION AND COMPARISON

We use the following four measurements to evaluate our MCA approach and compare its performance with the SMS algorithm discussed in [15]:

1) Total number of clusters $M$.
2) Total hop count, $H$, which is the summation of shortest hop distance between every node and its corresponding IBO over all the clusters.
3) Average hop distance $\bar{h}$ for a node, where $\bar{h} = H/(N-M)$.
4) Load variance $\text{var}(C)$ among all the clusters, where $\text{var}(C) = 1/M \cdot \sum_{i=1}^{M} (C_i - \bar{C})^2$ and $\bar{C} = N/M$.

First, we use the example in Drabu’s work [15] to compare the results between ours and theirs with $R_h = 3$ and $C_{\text{max}} = 8$. Fig. 4(a) and 4(b) shows the result using SMS algorithm and our proposed algorithm, respectively. The comparison in terms of four measurements obtained from Fig. 4 is listed in Table 1. From this table we can see that compared with SMS, our approach reduces the total hop count and average hop distance, though the number of clusters and load variance are the same. The load variances are zero since all the 40 nodes in the network are connected and the constructed clusters can satisfy all of the QoS constraints with the same cluster size equal to 8.

Now we compare the two algorithms’ performance in general case. We implement the two algorithms on 10 different network topologies with 50 nodes which are generated randomly with $R_h = 3$ and $C_{\text{max}} = 10$. The results are shown in Fig. 5, 6 and 7. The two algorithms give the same number of clusters if we relocate the cluster-head locations when implementing the SMS algorithm; if not, SMS algorithm will generate more clusters than ours. From these three figures we can see that, our MCA approach can get the better result in terms of three measurements, total hop count, average hop distance for a node, and load variance among clusters, compared to SMS algorithm. This is because: (1) we form clusters starting from the edge of the network and then go to the network center rather than start from the node with the maximum degree used in SMS; (2) we form the cluster based on certain criterions (illustrated in Procedure 2) and do not always add neighbor nodes to reach the maximum cluster size; (3) we update the IBO location at the final stage to minimize the average hop distance and balance the load.

**TABLE 1. MEASUREMENTS FROM FIG. 4**

<table>
<thead>
<tr>
<th></th>
<th>$M$</th>
<th>$H$</th>
<th>$\bar{h}$</th>
<th>$\text{var}(C)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMS</td>
<td>5</td>
<td>53</td>
<td>1.51</td>
<td>0</td>
</tr>
<tr>
<td>Proposed</td>
<td>5</td>
<td>50</td>
<td>1.43</td>
<td>0</td>
</tr>
</tbody>
</table>

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.
VI. CONCLUSION

In this paper, we have proposed the IBO placement algorithm in the hybrid EPON-WiMAX access networks while meeting several major QoS constraints. We have developed a modified clustering algorithm which can achieve the near-optimal solution, minimizing the number of clusters needed. The simulation results have been compared with the current literature and shown that our proposed algorithm generates at least the same number of clusters while reducing average transmission delay and link failure probability, balancing load for each cluster, and consequently improving the total network performance.

Future work can be investigated as IBO placement problem while take into account non-uniform user demand, channel interference, multi-radio and multi-channel routing algorithms, etc.

Acknowledgment

This work was partly supported by NSF under grant CNS 0619693.

REFERENCES