Passive Optical Network Layout in Manhattan

Samee Ullah Khan

***Abstract*— In this paper, we discuss the Passive Optical Network (PON) deployment on an arbitrary grid. We show that this problem in general is NP–hard. We propose an algorithm, which guarantees a solution of 3–approximation to the optimal deployment, and further argue that this is the best lower bound achievable in our case.**

***Index Terms*— Optimization methods, optical fiber communication, passive optical network, optical access network.**

I. Introduction

THE "last mile problem" is often the *bottleneck* in providing efficient and cost effective service to the end–users. Not only must the carries satisfy today's ever growing bandwidth demands by leveraging the limits of existing infrastructure, but also must plan for future subscriber needs. A new network infrastructure that allows more bandwidth, quick provision of services, and guaranteed QoS in a cost–effective and efficient manner is required [3].

The access lines that provide the users with a medium of communications have diversified in the recent times. High speed data communication services such as Asymmetric Digital Subscriber Line (ADSL) that uses metal wires, voice/ data communications services that use CATV networks, and wireless devices such as portable phones, are giving users/ vendors a wide range of media choices. PON has been considered the best choice for cost–effective delivery of high bandwidth data directly to the customers. Employing PON technology in Fiber To The Home (FTTH) architecture, makes quality, high–speed, and cost–effectiveness possible. Moreover coupling it with ATM technology, makes Quality of Service (QoS) answerable for users [2].

A basic architecture of PON is shown in Figure 1. The main component of PON is an optical splitter device. Depending on which direction the light is travelling, it splits the incoming light and distributes it to multiple fibers towards Optical Network Termination (ONT), or combines it into one towards Optical Line Terminal (OLT) [1]. The PON when included in the FTTH architecture, runs an optical fiber from the Central Office (CO) to an optical splitter and on into an ONT which then distributes it to the subscriber's location. The OLT is mostly located in the CO, or sometimes in an outside plant, or in a building. Fiber To The Cabinet (FTTCab) architecture runs an optical fiber from the CO to an optical splitter and then onto the Optical Network Unit (ONU), where the signal is converted to feed the subscribers over a twisted copper pair. The PON technology uses a double–star architecture. The first star topology centers at the OLT, and the second at the optical splitter.

S. U. Khan is with the Department of Computer Science and Engineering, The University of Texas at Arlington, TX–76019, USA (e–mail: sakhan@cse.uta.edu).

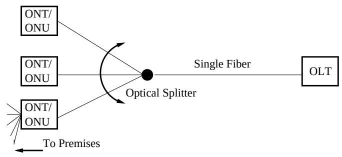

Fig. 1. Basic PON System Architecture

II. Problem Description

We can formulate the problem of optimal PON Network Layout (PNL) as a graph theoretical problem. Consider a graph $\mathcal{G}(V, E)$, such that V represent the physical locations of the subscriber's, CO, and another location acquired by the CO to expand its network, and E represent the communication lines between two V_i's. If there is no direct communication line $c(i, j)$ between V_i and V_j, we consider the shortest path between them measured in terms of simple distance or cost constraints. Without the loss of generality we assume that $c(i, j) = c(j, i)$. For simplicity we do not further sub–divide V into the obvious categories that represent the locations of OLT, ONT, CO, optical splitters and the subscribers. We can now formulate the PNL problem as follows:
"Given an undirected graph $\mathcal{G}$, find the locations of ONTs and splitters such that the cost of the equipment is minimized and for QoS the maximum distance from an ONT to a splitter and from a splitter to the OLT is minimized".

We assume that the OLT is residing inside the CO. The problem definition does not consider the optimization of ONT to the customer premises. This is due to the fact that the distance from ONT to the premises is negligibly small, and fault tolerance for a failed ONT can be answered by replacing the connection from a nearby ONT. We will treat ONT and ONU as more or less the same entity.

A. Preliminaries

In this section, we will define the terms and conditions necessary for deriving the proof of NP–hardness and our proposed algorithm.

Definition 1: A dominating set ($dom(\mathcal{G})$) of a graph $\mathcal{G}$ is a subset ($S \subset V$) such that every vertex in $V - S$ is adjacent to a vertex in S. Computing a $dom(\mathcal{G})$ is NP–hard.

Definition 2: A square of a graph $\mathcal{G}$, $\mathcal{G}^2$ is a graph containing an edge (u, v) whenever $\mathcal{G}$ has a path of length two between u and v, and $u \neq v$.

Definition 3: A star graph is a tree on $n + 1$ nodes with one node having vertex degree of n and the others having vertex degree 1.

Definition 4: An independent set (I) of a graph $\mathcal{G}$ is a subset of the vertices such that no two vertices in the subset represent an edge of $\mathcal{G}$. The cardinality of the independent set is the measure of the number of vertices included in the set, i.e. $|V|$.

Definition 5: The maximal independent set (M) of a graph is maximum cardinality of a set of vertices such that no two vertices in the subset represent an edge of $\mathcal{G}$.

B. PNL is Hard

Essentially our problem definition consists of two major optimization steps, i.e. ONT to splitters and splitters to the OLT. Showing the hardness of the problem over one optimization step would be sufficient to show that the over all the problem is hard. If we consider the optimization of the first phase of the problem, i.e. minimize the maximum distance of ONTs to the splitters and reduce the cost of the equipment, this can be solved by associating cost to vertices V (equipment cost) and edges E (fiber cost). Thus the problem reduces to finding the smallest number of minimum cost edges from a splitter to an ONT, such that the chosen set of edges identify vertices that connect in a min–max fashion.

Lemma 1: PNL problem is at least as hard as to find the minimum $dom(\mathcal{G})$.
Proof: Suppose that the edges are indexed in ascending order of associated cost, i.e. $cost_{e_1} \leq cost_{e_2} .... \leq cost_{e_m}$, than $dom(G)$ would denote the size of the minimum cardinality dominating set in $\mathcal{G}$. Thus solving PNL problem is no less easier than finding the $dom(\mathcal{G})$ [4]. ■

III. Our Approach

Here we describe our PNL algorithm. As mentioned in section II-B, the PNL problem is a two–step optimization problem. The first part deals with the location estimation of ONTs. This step we only concentrate on minimizing the fiber length and not the ONT equipment cost, since ONTs are not costly and they should be widely distributed in order to incorporate the random incursion and removal of subscribers. The second step takes as input the locations of ONT's which are used to optimize the splitter locations and reduce both the fiber and equipment cost. We assume that the edges of the graph $\mathcal{G}$ have the triangle inequality property. Let $s(u)$ denote the lightest neighbor of u in I. Let $S = \{s(u)|u \in M\}$. Thus, S represents the set of weighted vertices $w(v)$ such that once picked, they form a clique in graph $\mathcal{G}$. The input to the PNL algorithm would be simple the graph $\mathcal{G}$ and k which is the number of clustered population.

Lemma 2: The PNL algorithm is complete and will identify a solution, if there exists one.
Proof: The first part of the algorithm lines 1–5 (optimizing ONT locations), is based on the fact that $|I| \leq dom(\mathcal{G})$. Let D be the minimum dominating set in a graph $\mathcal{G}$. Then, in the worst case $\mathcal{G}$ contains $|D|$ stars spanning all vertices. Since each of these stars would form a clique in $\mathcal{G}^2$, therefore, $\mathcal{G}^2$ has at least $|D|$ cliques spanning all the vertices. Thus, one can pick at most one and only vertex from each clique. The second part of the algorithm lines 6–12 (optimizing ONT to splitter), works the same way as the first part except that now in the weighted graph the neighbors are picked in $\mathcal{G}$ and not in $\mathcal{G}^2$. ■

PNL Algorithm
Input: $\mathcal{G}$ and k.
Output: $\mathcal{G}'$ (Final PNL).

1. Construct $\mathcal{G}_1^2, \mathcal{G}_2^2, ..., \mathcal{G}_m^2$.
2. Compute M_i in each $\mathcal{G}_i^2$.
3. Find the smallest index i so that $|M_i| \leq k$, say M_j.
4. Construct back $\mathcal{G}_j$ from M_j.
5. Input $\mathcal{G}_j$ and $|M_j|$ for step 5.
6. Construct $\mathcal{G}_{1j}^2, \mathcal{G}_{2j}^2, ..., \mathcal{G}_{mj}^2$.
7. Compute M_{ij} in each $\mathcal{G}_{ij}^2$.
8. Compute $S_i = \{s_i(u)|u \in M_{ij}\}$.
9. Find the minimum index i so that $w(S_i) \leq w(D)$, say S_j.
10. Return S_j.
11. Compute $\mathcal{G}' = \min_{M_{ij}} \max_{s_i(u)} \sum_{i=0}^{|S_j|}$.
12. Return $\mathcal{G}'$.

Lemma 3: If a solution is identifiable, the PNL algorithm computes the solution in $\mathcal{O}(n^5)$.
Proof: The bottleneck is computing the square of the subgraphs in steps 1 and 6 computable in $\mathcal{O}(n^3)$, that are used to identify the maximal independent set. Maximal independent set can be identified in $\mathcal{O}(n^2)$. Since these two operations are interdependent, we cannot simply parallelize them. Thus, the overall running time is $\mathcal{O}(n^5)$. ■
Although the running time of $\mathcal{O}(n^5)$ is high, we argue that PNL algorithm will be used at initial layout of the network. If so, it is worth spending time to optimally and cost-effectively deploy a passive optical network.

Lemma 4: The PNL algorithm has a lower bound of 3–approximation to the optimal algorithm.
Proof: The key observation is that I in a graph also forms the $dom(G)$. If a vertex v is not dominated by I, then $I \cup \{v\}$ would form I, contradicting definition 3. Thus, there exist starts (Definition 3) in $\mathcal{G}_j^2$, such that their centers lay on M_j and covering all the vertices. By the assumption that triangle inequality holds, these stars would use edges of cost at most $2 \cdot cost(e_j)$. Moreover each star center would be adjacent to a vertex in S_j, using an edge of cost at most $2 \cdot cost(e_j)$. Thus, forcing to move each center to the adjacent vertex in S_j, and in turn redefining the stars. Since the edge costs exhibit triangle inequality, the largest edge cost in constructing the stars is at most $3 \cdot cost(e_j)$. ■

IV. Experiment and Discussion

We evaluated our proposed algorithm, from the point of view of saving the fiber and equipment cost. We also identified the relationship between the splitters, ONT's and the fiber length being used during the experiment run. We considered a 10 sq. miles *Manhattan* grid with 50m. apart streets. The choice of Manhattan grid was due to the fact that PNL algorithm will fail if the layout is performed on an arbitrary graph whose edges do not show triangle inequality (Definition 3). The locale was populated with 20K personnel (random distribution), clustered into small, medium and large businesses. We assumed that small business has on average 6 personnel, medium has 60 and large franchise has 2000 employees on average. The assumed simulated traffic per person per business type can be seen in table I. Fiber cost was assumed to be $100/m, ONT at $50/unit, splitter at $600/unit and OLT at $250/unit. To evaluate our solution quality, we

TABLE I
TRAFFIC SIMULATION PATTERNS

Type	**Traffic**		
	Small Business	**Medium Business**	**Large Business**
Voice	–	–	–
Web	0.9MB	1.35MB	1.8MB
E–mail	15kB	150kB	500kB
FTP	5MB	7MB	7MB
Fax	4kB	4.8kB	4.8kB
Total/person	5.919MB	8.505MB	9.3045MB
Total/busines	35.514MB	510.3MB	18609MB

first generated a random network layout. The main purpose was to compare PNL algorithm with arbitrary layouts. The parameters such as the number of ONTs, splitters, OLT and the fiber used were observed. In the second phase we ran our proposed algorithm to generate the network layout. The difference (drop/ increase) in solution quality was observed. For the evaluation of each parameter, i.e. maximum allowable fiber, ONT deployment distance etc., we generate a total of 5 different random network layouts. The average fiber used, ONTs and splitters were recorded in all the 5 runs, to have a comprehensive comparison with our proposed algorithm.

In the first series of experiments we evaluated the general pattern observed by our proposed algorithm. Figure 2, shows

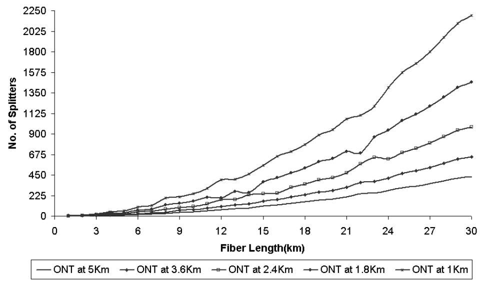

Fig. 2. ONT Distance vs. Fiber Length.

the relationship between the fiber length and the number of splitters used. The number of splitters is directly related to the splitting number. Thus, more splitters would guarantee higher branching factor in the network and which in turn would guarantee extra fiber usage. The maximum distance allowable for the placement of ONT was varied from 1km to 5km. A consistence increase was observed with the decrease in the allowable ONT distances. The next two experiments (Figures 3–4) were used to evaluate the fiber and number of splitters being saved as compared to that of the randomized network layout placement. For each run, the average of 5 random layouts was compared with PNL algorithm. On average the PNL algorithm saved 14–59% splitters being used and 13–64% fiber dependent on the constraints on the maximum allowable fiber and maximum allowable distance of ONT placements respectively.

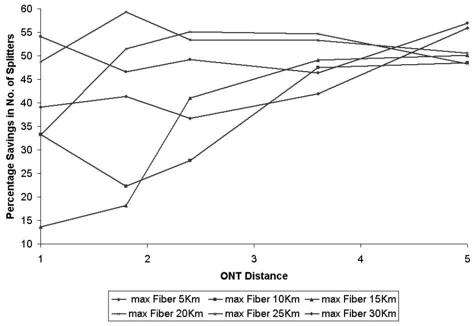

Fig. 3. Allowable Fiber vs. No. of Splitters Saved.

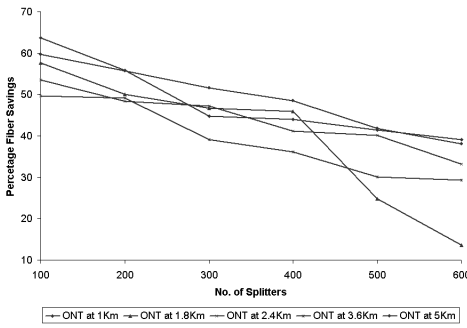

Fig. 4. No. of Splitters vs. Fiber Saved.

V. CONCLUSIONS

We proposed a 3–optimal algorithm for the generalized Passive Optical Network Layout (PNL) problem. The experimental results revealed significant saves in fiber, ONT and reduce branching factor as compared to random layouts.

REFERENCES

[1] C.-J. Chae, S.-T. Lee, G.-Y. Kim, and H. Park. A PON system suitable for internetworking optical network units using fiber bragg grating on the feeder fiber. *IEEE Photonics Technology Letters*, 11:1686–1688, December 1999.

[2] F. J. Effenberger, H. Ichibangase, and H. Yamashita. Advances in broadband passive optical networking technologies. *IEEE Communications*, 39(12):118–124, 2001.

[3] D. Kettler, H. Kafka, and D. Spears. Driving fiber to the home. *IEEE Communications*, 38(11):106–110, 2000.

[4] S. U. Khan. Why should we pay more for layout designers? In M. I. Khan and A. A. S. Awwal, editors, *Photonic Devices and Algorithms for Computing V*, volume 5201. SPIE, August 2003.