Optimal Strategy for Graceful Network Upgrade

Ram Keralapura, Chen-Nee Chuah, Yueyue Fan
University of California, Davis
{rkeralapura, chuah, yyfan}@ucdavis.edu

ABSTRACT

One of the critical aspects of network management that has not received much attention is network upgrade. This paper addresses the question of "how to add new nodes and links into an operational network in a graceful manner so that the perceived network performance from the perspective of existing customers does not deteriorate?". We propose a two-phase framework to find the optimal upgrade strategy: first, deciding what nodes should be added and how they should be connected to existing topology, and second, deciding the ideal sequence to add these new nodes and links. We formulate the first phase as a non-linear optimization problem and the second phase as a multistage dynamic programming problem. Through a numerical example, we show the feasibility of this framework and demonstrate the advantages of our multistage approach in determining an ideal upgrade sequence. The results also highlight the significance of incorporating network performance (for ex, service availability) into the two-phase framework to achieve minimal impact to existing customers.

Keywords

Graceful Network Upgrade, Multi-Stage Network Upgrade

1. INTRODUCTION

Internet service providers (ISPs) face a plethora of critical challenges in the area of network management. One such challenge is the problem of *graceful network upgrade*, where ISPs need to add new nodes and links into their operational network in a graceful manner so that the perceived network performance from the perspective of existing customers does not deteriorate. Although graceful network upgrade is an important aspect of network management, it has not received much attention in the past. To the best of our knowledge, as of today, there are no publicly available studies that provide a framework for adding new nodes and links into existing operational networks. Most of the previous works have focused on upgrade in the context of either upgrading individual network components [10] or capacity provisioning [4].

From the perspective of ISPs, planning a successful network upgrade involves three main steps: (i) identifying a set of potential locations where new nodes can be added, (ii) determining a subset of all the identified locations where nodes should be added (along with the links) such that it results in maximum revenue for the ISP given certain budget and performance constraints, and (iii) identifying an ideal sequence for adding nodes and links into the network such that the upgrade process has minimum impact on the existing customers. There are several techniques like [6, 2, 9, 3], that can be used to address the first step from an economic viewpoint. The focus of our work is on the last two steps assuming that an ISP has already identified a set of potential locations for adding new nodes.

In this paper, we propose a *two-phase multistage* framework to determine the optimal network upgrade strategy. One of the key features of this framework is the emphasis on considering network performance not only from the perspective of the ISP but also from the perspective of its customers at every stage. In *Phase-1*, we determine the nodes and links that should be added to the network by formulating the problem as a non-linear optimization problem. In *Phase-2*, we determine an ideal sequence for adding the new nodes and links to ensure minimal impact on network performance using multistage dynamic programming (DP). The time period between different stages in this phase is typically in the order of several months to few years. We believe that multistage node addition in Phase-2 is critical for an ISP for several reasons: (i) It would be extremely hard to debug problems that arise if all the identified nodes and links are added simultaneously. Given that the network is operational during the upgrade process, it is important to ensure that the customers do not suffer from severe performance degradation; (ii) ISP may face delays due to node/link construction schedule and certain nodes/links could be available sooner than others; (iii) The budget may only become available sequentially.

The main contributions of our work are:

- We propose a two-phase framework that can be used to determine the optimal network upgrade strategy in ISP backbone networks. We believe that with only minor modifications, this framework can be applied in the context of different network service providers like cellular service providers, WiMax service providers, etc.
- We formally define and formulate the problem in Phase-1 (optimal end state determination phase) as a non-linear optimization problem and Phase-2 (multistage node addition phase) as a multistage DP problem. While the problem in Phase-1 can be solved using well established non-linear programming techniques (such as simulated annealing, genetic algorithms, and tabu search), the problem in Phase-2 can be solved using multistage DP technique.
- In both the phases of our framework, we highlight the importance of considering network performance from the perspective of both customers and ISPs. We show that such an approach performs better than the traditional approach where network performance is considered only from the perspective of an ISP.

SIGCOMM'06 Workshops September 11-15, 2006, Pisa, Italy.

- We show the feasibility of our framework using a numerical example. The results clearly demonstrate the significance of choosing the correct sequence of nodes to add into the network.

2. GRACEFUL NETWORK UPGRADE

Our focus in this work is on tier-1 ISP backbone networks where nodes correspond to Points-of-Presence (PoPs) and links correspond to IP-layer virtual connections on top of the underlying optical fiber. Consider a tier-1 ISP backbone network that spans the continental US. All the nodes in such a network are strategically placed in locations such that the ISP can serve a large number of customers who can easily connect to its nodes, thus resulting in the maximum possible revenue. Although such a network could be ideal today, it is important for ISPs to add more nodes into their backbone network to serve customers in regions that have the potential to become business/residential hubs in the future.

In the context of our work, network upgrade *does not* refer to upgrading individual components in the network. An obvious approach for an ISP is to add nodes at all possible locations and connect them in a full mesh to give the best performance, but such a strategy is usually infeasible due to budget constraints. We assume that the entire upgrade procedure has a maximum fixed budget, b, that cannot be exceeded. We also assume that an ISP:

- can reliably estimate the future traffic demands to identify all potential physical locations where it is beneficial to install new nodes [2, 9, 3].
- can accurately estimate the revenue/profit associated with the addition of new nodes in the locations identified above [6]. Each node addition results in a revenue/profit that is independent of other node additions in the network.

2.1 Network Performance from a Customer's Perspective

Recent studies in tier-1 carriers have found that performance deterioration inside tier-1 ISP networks are mainly caused by *link failures* that occur frequently due to several reasons like fiber cuts, software bugs, and hardware errors [5]. The authors in [8] developed a general model that captures the duration and frequency of occurrence of such link failures. Based on this link failure model and the observed transient behavior during routing convergence, we proposed two metrics that capture the network service availability from both customer's and ISP's perspective [7]:

- ***Service Disruption Time* (SD Time)**: This metric captures the average time for which a customer connected to an ISP through ingress node gets disconnected from various prefixes in the Internet due to link failures. It is important for an ISP to ensure that the SD time for all the nodes in their network is below the acceptable values specified in their service level agreements (SLA).
- ***Traffic Disruption* (TD)**: This metric captures the total traffic disrupted between different nodes in the network due to link failures. TD should be included in the 'cost' function since an ISP typically has to compensate customers when the total TD for the customer exceeds the threshold value specified in the SLA.

Using extensive simulations, we previously showed that SD time and TD capture several important performance characteristics of an operational network. We also showed that, given operational network conditions, these metrics can be statically computed (no active or passive measurements necessary) using the algorithm in [7]. In the rest of this paper, we use SD time and TD (either as constraints or as high cost penalties when violated) to capture the network performance from a customer's perspective.

Note that SD time and TD metrics depend on operational network conditions like traffic demand, exit points for various Internet prefixes (i.e., network peering points), and shortest paths between different nodes in the network [7]. Adding new nodes and links could change these operational network conditions, for example, adding new nodes could change the network traffic matrix due to additional traffic sources and sinks, adding new links could change the shortest paths between different nodes, and adding peering links could change the exit points for various prefixes. These factor can affect the network performance as seen by customers connected to the network. Our goal in this work is to minimize the performance deterioration due to changing operational network conditions. It is also important to note that SD time and TD metrics can be easily replaced by other context-specific metrics while keeping the same overall framework proposed in this paper.

2.2 Problem Description

Consider a network with n nodes and l links. We use $\mathcal{N}$ to represent the set of all nodes and $\mathcal{L}$ to represent the set of all links in the original network. We use several attributes to capture different characteristics of the network:

- **Adjacency Matrix, $\mathbf{M}$:** An $n \times n$ matrix where each element is either a 1 or 0. $M_{i,j} = 1$ indicates that there is a link that directly connects node i and node j, and $M_{i,j} = 0$ indicates that there is no link between i and j.
- **Distance Matrix, $\mathbf{D}$:** An $n \times n$ matrix indicating the physical distance between different nodes in the network.
- **Service Disruption Time Matrix, $\mathbf{S}$:** An $l \times n \times n$ matrix where $S_{i,j,k}$ represents the total service disruption time for the source-destination pair $j - k$ due to failure of link i.
- **Traffic Disruption Matrix, $\mathbf{T}$:** Similar to S this is a $l \times n \times n$ matrix where $T_{i,j,k}$ represents the total traffic disrupted between the source-destination pair $j - k$ due to failure of link i.

Note that the adjacency matrix, $\mathbf{M}$, and the distance matrix, $\mathbf{D}$ are symmetric matrices since we assume that all the links are bidirectional, while service disruption time matrix, $\mathbf{S}$, and traffic disruption matrix, $\mathbf{T}$, may not be symmetric.

Let g represent the total number of locations where nodes can be placed to upgrade the network. We represent the revenue/profit from an ISP's perspective by $\mathbf{r}$ ($g \times 1$ vector) where each element r_k indicates the revenue/profit when the new node at position k is added to the network. Similarly we use $\mathbf{nc}$ ($g \times 1$ vector) to represent the estimated building and installation cost for new nodes at different potential locations. Note that the cost of building and installing nodes at different locations could be different. We use lc to indicate the cost per unit length of installing new links. The cost here refers to the cost of installing optical fiber between different nodes. Finally we assume that all the links (in the current network and upgraded network) have enough capacity to carry the traffic demands and hence we ignore using link capacities as variables in our problem formulation. However it is easy to include link capacities as variables in our two-phase network upgrade framework.

The graceful network upgrade problem can be defined as: Given a total budget constraint, find the locations where nodes should be added and how they should be connected to the other nodes (original and newly added nodes) in the network so as to maximize the revenue from the network while maintaining certain performance constraints. We should also determine the optimal sequence to perform this upgrade. During the network upgrade process we should ensure that the network performance (in terms of service disruption time and traffic disruption) from the perspective of customers connected to the original network does not deteriorate.

Let $\mathbf{k}'$ ($(n+g) \times 1$ vector) represent a 0/1 vector where ${k'}_i = 1$ indicates that a node exists at location i, and ${k'}_i = 0$ indicates that no node exists at location i. Since we do not remove any of the

Symbol	Description
n	Number of nodes in original network
l	Number of links in original network
$\mathcal{N}$	Set of all nodes in original network
$\mathcal{L}$	Set of all links in original network
$\mathcal{N}'$	Set of all nodes in upgraded network
$\mathcal{L}'$	Set of all links in upgraded network
$\mathbf{M}$ $[n \times n]$	Adjacency matrix of original network
$\mathbf{D}$ $[n \times n]$	Distance matrix of original network
$\mathbf{S}$ $[l \times n \times n]$	SD time matrix of original network
$\mathbf{T}$ $[l \times n \times n]$	TD matrix of original network
α	Link utilization threshold
b	Total budget for network upgrade
g	Number of potential locations to add nodes
$\mathbf{r}$ $[g \times 1]$	Revenue vector for potential node locations
$\mathbf{nc}$ $[g \times 1]$	Node installation cost
lc	Link installation cost per unit length
$\mathbf{k}'$ $[(n+g) \times 1]$	Node placement vector for upgraded network
$\mathbf{M}'$ $[(n+g) \times (n+g)]$	Adjacency matrix of upgraded network
$\mathbf{D}'$ $[(n+g) \times (n+g)]$	Distance matrix of upgraded network
$\mathbf{u_n}$ $[n \times 1]$	Unit matrix of size $n \times 1$

Table 1: Notations used in problem formulation

original nodes in the network during the upgrade process, we have

$$\mathbf{k}' = \begin{pmatrix} \mathbf{u_n} \\ \mathbf{k}'' \end{pmatrix} \qquad (1)$$

where $\mathbf{u_n}$ ($n \times 1$ vector) is a unit vector with all its elements equal to 1, and $\mathbf{k}''$ is a $g \times 1$ vector whose elements can be 0/1 depending on the locations where the nodes are added to the network.

Let $\mathbf{M}'$ represent the adjacency matrix of the upgraded network.

$$\mathbf{M}' = \begin{pmatrix} \mathbf{M} & (\mathbf{M}'')^T \\ \mathbf{M}'' & \mathbf{M}''' \end{pmatrix}$$

where $\mathbf{M}$ is the adjacency matrix of the original network; $\mathbf{M}''$ is a $g \times n$ matrix that indicates the links between the nodes added during the upgrade process and the original nodes in the network; $\mathbf{M}'''$ is a $g \times g$ matrix that indicates the connectivity between the new nodes added into the network during the upgrade process.

In order to determine optimal strategy for network upgrade, we have to determine the values of $\mathbf{k}''$, $\mathbf{M}''$, and $\mathbf{M}'''$ such that all the constraints are met. We will discuss more about the constraints in Section 3. A summary of all the notations used in this paper is presented in Table 1.

3. PROBLEM FORMULATION

We decouple the graceful network upgrade problem into the following two phases:

3.1 Phase-1: Finding the Optimal End State

The main objective of this phase is to determine an end state that maximizes the total revenue/profit for the ISP. This can be formulated as follows:

Objective function: Maximize $((\mathbf{k}'')^T.\mathbf{r})$ subject to the following constraints:

Budget constraint: The total cost of building and installing the nodes and links should be less than the total budget.

$$\frac{lc}{2}\left\{\mathbf{u}_{n+g}^T.(\mathbf{D}' \otimes \mathbf{M}').\mathbf{u}_{n+g} - \mathbf{u}_n^T.(\mathbf{D} \otimes \mathbf{M}).\mathbf{u}_n\right\} + ((\mathbf{k}'')^T.\mathbf{nc}) \leq b$$

where $\otimes$ operator represents element-wise multiplication of the two matrices and $\mathbf{D}'$ is the distance matrix of the upgraded network. Since all the links are bidirectional and are considered twice in the adjacency and distance matrices we need the factor 2 in the denominator of the first term in the above equation.

Node degree constraint: One of the basic requirements for resiliency that all ISPs require is that every node in the network has at least two independent paths to route traffic originating at the node. In other words, the degree of all the nodes added into the network should be at least 2.

$$\sum_{j=1}^{n+g} \mathbf{M}'_{i,j} \geq 2.\mathbf{k}_i \quad \forall i = n+1, n+2, ..., n+g$$

SD time constraint: For an ISP it is important to ensure that the network performance experienced by customers connected to the nodes in the original network does not deteriorate. Let $\mathbf{S}'$, $\mathcal{N}'$, and $\mathcal{L}'$ represent the SD time matrix, set of nodes, and set of links respectively in the final upgraded network. Hence we can write the SD time constraint as,

$$\sum_{i \in \mathcal{L}', q \in \mathcal{N}'} \mathbf{S}'_{i,j,q} \leq \sum_{i \in \mathcal{L}, q \in \mathcal{N}} \mathbf{S}_{i,j,q} \quad \forall j \in \mathcal{N} \qquad (2)$$

TD constraint: Similar to SD time, the traffic disruption experienced by customers should not become worse after an upgrade. We use $\mathbf{T}'$ to represent the TD matrix of the upgraded network. Hence,

$$\sum_{i \in \mathcal{L}', q \in \mathcal{N}'} \mathbf{T}'_{i,j,q} \leq \sum_{i \in \mathcal{L}, q \in \mathcal{N}} \mathbf{T}_{i,j,q} \quad \forall j \in \mathcal{N} \qquad (3)$$

Link utilization constraint: Maximum link utilization has frequently been used by ISPs as a measure of network quality and failure resiliency [1]. Network operators always want to ensure that the maximum link utilization in their network is below a certain threshold value (say α). We represent the utilization of a link i by LU_i. Hence,

$$\max_{\forall i \in \mathcal{L}'} LU_i \leq \alpha$$

The linearity or non-linearity of Phase-1 depends mainly on the performance metrics that are chosen as constraints. The above formulation is non-linear since $\mathbf{S}$ and $\mathbf{T}$ are non-linear terms [7] (details not given due to space limitation). This problem can be solved by well-established non-linear programming techniques such as simulated annealing, genetic algorithms, and tabu search.

Note that, in the above formulation: (i) Absolute (not relative as in Eqns 2 and 3) values for SD time and TD constraints can be used, i.e., SD time and TD greater than particular threshold values (based on the SLA between ISP and customers). (ii) Constraints can be imposed on SD time and TD experienced by *new nodes* added to the network depending on the SLA that the ISP plans for the customers connecting to the new nodes. However, we currently ignore these two conditions in both Phase-1 and Phase-2, since our focus in this paper is to establish the usefulness of the basic framework without adding secondary details.

3.2 Phase-2: Multistage Node Addition

The main objective of this phase is to determine the optimal sequence in which nodes and links should be added into the network. In this phase, we make two assumptions:

- We assume that we can obtain a solution from the first phase, i.e., the values of $\mathbf{k}'$ and $\mathbf{M}'$. In other words, we have complete knowledge of the initial network state that we start with and the final state that we intend to reach.
- We also assume that the ISP wants to add no more than one node at a time. Hence every stage in our multistage node addition strategy adds one node and all the links associated with that node into the network.

We formulate Phase-2 as a multistage DP problem, and use network performance at every stage as the cost involved in upgrading the network from one stage to the next. We characterize the network performance by four metrics: service disruption time, traffic disruption, max link utilization, and the degree of different nodes in the network. Note that these metrics can be used either as constraints or as costs in the utility function. We choose to use these metrics as costs since using them as constraints may not result in any feasible solution. On the other hand, assigning a high cost when the performance criteria is not met will lead to feasible solutions albeit higher cost. In this work, we treat all the four performance criteria with the same importance. However, a weighting scheme can be introduced to distinguish their importance.

We use the following notations in the dynamic program problem formulation. Let $\mathcal{N}^e$ and $\mathcal{L}^e$ represent the set of nodes and links respectively in the network after e node addition stages. Hence, $|\mathcal{N}^e| = |\mathcal{N}| + e = n + e$. We use $\mathbf{v}_n^e$ to represent a 0/1 vector of nodes that are remaining to be added after e stages. Let h be the node added to the network in stage e. We can represent the optimal return function after stage e by $\Phi^e(\mathbf{v}_n^e, h)$. Note that the return function after any stage depends on the location of the node added and the locations of the nodes remaining to be added. Finally, we represent SD time and TD values after stage e by $\mathbf{S}^e$ and $\mathbf{T}^e$.

We model the cost associated with network performance at stage e due to the addition of node h as follows:

- **Cost of SD time, $c_{SD}^{e,h}$:** ISPs require the SD time experienced by customers (i.e. nodes) in the original network should not increase after any stage of network upgrade. Hence we define a cost function such that the cost is very large only when the current SD time for the original nodes is larger than the original SD time.

For every node, i, in the original network we define the cost of SD time as:

$$c_{SD}^{e,h}(i) = \left\{ \frac{2\delta^{e,h}(i) - \delta(i)}{\delta(i)} \right\}^{2\omega_{SD}(i)} \quad \forall i \in \mathcal{N} \tag{4}$$

where $\omega_{SD}(i)$ is a high order index value that ensures that the cost is high only when $\delta^{e,h}(i) > \delta(i)$ and the cost is very small otherwise. Also,

$$\begin{aligned} \delta^{e,h}(i) &= \sum_{j \in \mathcal{L}^e, q \in \mathcal{N}^e} \mathbf{S}^e{}_{j,i,q} \\ \delta(i) &= \sum_{j \in \mathcal{L}, q \in \mathcal{N}} \mathbf{S}_{j,i,q} \end{aligned}$$

Hence the overall cost of SD time in the network is,

$$c_{SD}^{e,h} = \sum_{i \in \mathcal{N}} a_{SD}(i).c_{SD}^{e,h}(i) \tag{5}$$

where $a_{SD}(i)$ is a weighting factor that indicates the importance of the performance of node i to the ISP.

- **Cost of TD, $c_{TD}^{e,h}$:** We can express the cost of TD in a very similar fashion as cost of SD time. Hence,

$$c_{TD}^{e,h}(i) = \left\{ \frac{2\tau^{e,h}(i) - \tau(i)}{\tau(i)} \right\}^{2\omega_{TD}(i)} \quad \forall i \in \mathcal{N} \tag{6}$$

where $\omega_{TD}(i)$ is a high order index value and,

$$\begin{aligned} \tau^{e,h}(i) &= \sum_{j \in \mathcal{L}^e, q \in \mathcal{N}^e} \mathbf{T}^e{}_{j,i,q} \\ \tau(i) &= \sum_{j \in \mathcal{L}, q \in \mathcal{N}} \mathbf{T}_{j,i,q} \end{aligned}$$

Hence the overall cost of TD in the network is,

$$c_{TD}^{e,h} = \sum_{i \in \mathcal{N}} a_{TD}(i).c_{TD}^{e,h}(i) \tag{7}$$

where $a_{TD}(i)$ is a weighting factor that indicates the importance of the performance of node i to the ISP.

- **Cost of node degree, $c_{degree}^{e,h}$:** Since the most common failures in networks are single link failures, it is important to ensure that all the nodes have a node degree of at least 2. Note that during the multistage upgrade process it may not be feasible to ensure that the newly added node will have a degree of at least 2. However after the completion of the upgrade process all the nodes will have a degree of at least 2 since our solution in Phase-1 guarantees this. Hence a high cost at any stage resulting in degree of 0 or 1 will help in finding a strategy with a node degree of at least 2 in all the stages.

$$c_{degree}^{e,h}(i) = \left\{ \frac{1}{(\kappa^{e,h}(i))(\kappa^{e,h}(i) - 0.5)} \right\}^{2\omega_{degree}(i)} \quad \forall i \in \mathcal{N}^e \tag{8}$$

where $\omega_{degree}(i)$ is a high order index value to ensure that the cost for a node with small degree is high, and $\kappa^{e,h}(i)$ is the degree of node i after the addition of node in location h in stage e. Hence the total cost for node degree for the network at stage e is:

$$c_{degree}^{e,h} = \sum_{i \in \mathcal{N}^e} a_{degree}(i) c_{degree}^{e,h}(i) \tag{9}$$

where $a_{degree}(i)$ is a weighting factor that indicates the importance of the performance of node i to the ISP.

- **Cost of link utilization, $c_{util}^{e,h}$:** A high utilization on a link implies that it will not be able to accommodate more traffic when there is a failure elsewhere in the network. Hence the operators set a threshold value (α) for the maximum link utilization.

We can express the cost of link utilization as:

$$c_{util}^{e,h} = \left\{ \frac{2\sigma^{e,h} - \alpha}{\alpha} \right\}^{2\omega_{util}} \tag{10}$$

where $\sigma^{e,h} = \max_{\forall i \in \mathcal{L}^e} LU_i$.

The total cost of upgrade at stage e is $c^{e,h} = c_{SD}^{e,h} + c_{TD}^{e,h} + c_{degree}^{e,h} + c_{util}^{e,h}$. The functional equation of the multistage dynamic programming formulation now becomes:

$$\Phi^e(\mathbf{v}_n^e, h) = \max_{\forall h \in \mathbf{v}_n^{e-1}} \left\{ -c^{e,h} + \gamma \Phi^{e+1}(\mathbf{v}_n^{e+1}, h_1) \right\} \tag{11}$$

with the boundary condition,

$$\Phi^{n'-n+1}(\mathbf{v}_n^{n'-n+1}, h_2) = 0 \tag{12}$$

where $h_1 \in \mathbf{v}_n^e$, $h_2 \in \emptyset$, and $n' = |\mathcal{N}'|$. Also, $\gamma \leq 1$ represents the discounting factor for future benefits. In the rest of this paper we use $\gamma = 1$.

4. NUMERICAL EXAMPLE

In this section, we illustrate the feasibility and advantages of using our two-phase multistage upgrade process described earlier. For the ease of exposition we use a simple example, although the framework can apply to real-world topologies of O(100) nodes.

We consider the topology in Figure 1 where nodes are situated in geographically distributed locations in the US. Our aim is to create a scenario similar to that of tier-1 IP backbone (PoP-level) topologies that span the continental US. The original topology that we

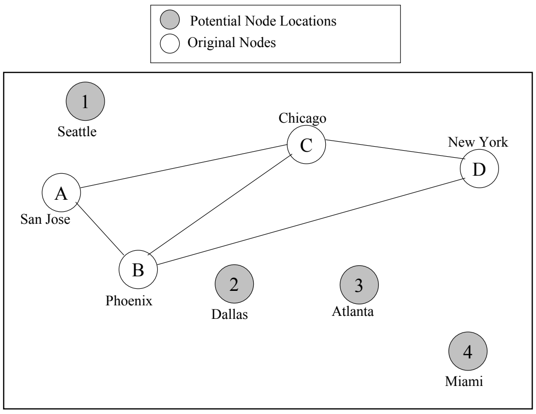

Figure 1: An example topology and potential node locations for new node placement

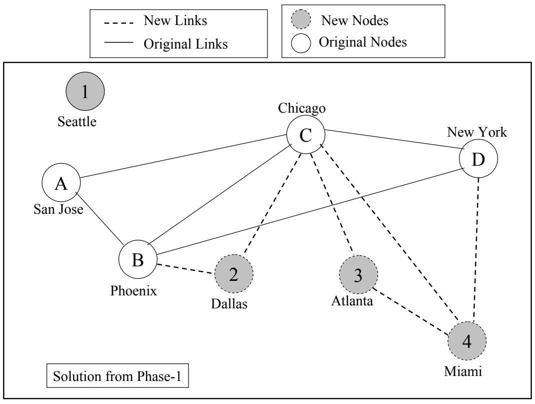

Figure 2: Final solution from solving the non-linear program in Phase-1 with specific network and budget constraints

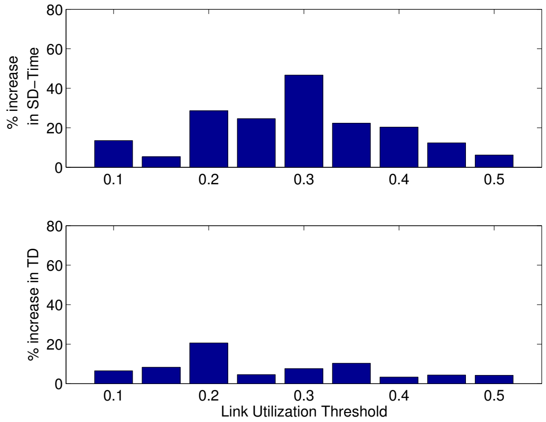

Figure 3: Impact on Service Disruption time and Traffic Disruption when using maximum link utilization and node degree as metrics to estimate network performance.

consider has four nodes (`A`, `B`, `C`, and `D`) and five links between them. Although IP-links are virtual links on top of the optical infrastructure, in this work we assume that the propagation delay for different links is dependent on the geographical distance between the nodes that they connect. Based on the findings in [7], we categorize the nodes in the network as large, medium and, small depending on the amount of traffic they generate (typically in the ratio 4:2:1). In our example we consider nodes `C` and `D` as large nodes, `B` as a medium node, and `A` as a small node. We generate our traffic matrix based on this assumption. Another important consideration in the network is the BGP prefix distribution among different nodes, i.e., BGP exit points to reach different prefixes in the Internet. SD time and TD for different nodes depend on the prefix distribution in the network [7]. We model the prefix distribution to incorporate the elephant and mice phenomenon where 80% of the traffic is destined to 20% of the prefixes. Finally we assume equal link weights for all the links in the network (i.e., minimum hop routing).

As a starting point we identify four distinct locations (`1`, `2`, `3`, and `4`) where new nodes can be added to the original network (see Figure 1). We assume that nodes `1` and `4` are large nodes, `3` is a medium node, and `2` is a small node. We also assume that an ISP can identify the potential locations for adding nodes and estimate the traffic and revenue for each of the locations. Note that these estimations could come with a number of uncertainties and dependencies among each other. However, as a first step, we assume that the traffic and the associated revenue generated by a node is independent of other nodes in the network. We defer the study of uncertainties in estimation and how we can design an adaptive approach for network upgrade to future work.

In our simple example we assume: $\mathbf{r} = \{45000, 25000, 38000, 50000\}$, $\mathbf{nc} = \{14000, 13000, 13800, 14000\}$, $lc = 1$, and $b = 50000$. Our first objective (i.e. Phase-1) is to determine which of the four potential nodes should be added to the network to maximize revenue and how should they be linked to the other nodes while ensuring that the budget, SD time, TD, node degree, and link utilization constraints are satisfied. We use the Tabu Search technique to explore the final solution. Since the problem is small we also use a brute force approach to confirm that the above solution is in fact the global optimal solution. The final solution that we obtained for Phase-1 is shown in Figure 2. We can see that only three of the four potential nodes were included.

Typically network operators use maximum link utilization in the network as a measure of network performance. However, link utilization does not consider the network performance from users' perspective and hence optimization decision based solely on link utilization could result in performance degradation for some/all customers. To reinforce this concept we find the solution to the problem in Phase-1 by ignoring the SD time and TD constraints, and use only link utilization and node degree as constraints for network performance. This approach will try to find a solution (referred to as *utilization solution*) that minimizes the maximum link utilization in the network among all the feasible solutions that maximize the revenue.

In Figure 3 we compare the SD time and TD values for nodes `A`, `B`, `C`, and `D` before and after the upgrade (using the utilization solution). We compute the utilization solution by varying α (the maximum link utilization threshold value) between 0.1 and 0.5, a range used in real world networks. We can see that, in the case of our simple network, SD time could increase up to 46% and TD could increase up to 20%. These values could vary significantly depending on the operating network conditions. The main conclusion from Figure 3 is that it is important to consider SD time and TD constraints while performing network upgrade.

Now that we have determined the optimal end state, our next objective is to determine an optimal strategy for multistage node

All Possible Solutions	*Stage-{1,2,3}*	Total Cost
Solution-1	$2,3,4$	1056.76
Solution-2	$2,4,3$	16.74
Solution-3	$3,2,4$	1090.33
Solution-4	$3,4,2$	1065.42
Solution-5	$4,2,3$	1.654
Solution-6	$4,3,2$	5.503

Table 2: All possible upgrade solutions for Phase-2 (Note: $\omega_{SD} = \omega_{TD} = \omega_{util} = 2$ **and** $\omega_{degree} = 5$)

Multistage Solution	*With SD time and TD*		*Without SD time and TD*	
	Node Added	*Links Added*	*Node Added*	*Links Added*
Stage-1	4	$4-C, 4-D$	4	$4-C, 4-D$
Stage-2	2	$2-B, 2-C$	3	$3-4, 3-C$
Stage-3	3	$3-4, 3-C$	2	$2-B, 2-C$

Table 3: Multi-stage network upgrade solution for Phase-2 with and without using SD time and TD as metrics for network performance.

addition. Here we assume that the ISP adds one node at a time, along with all the links associated with that node. We implemented the dynamic programming solution for Phase-2 using Matlab that takes the initial and final network conditions as input, and outputs the optimal sequence to reach the final stage from the initial stage.

In our current example, we need to add three nodes and six links into the network. Hence we need two intermediate stages to reach the final stage. Note that there are 6 different ways in which this can be accomplished. All possible solutions for multistage network upgrade are shown in Table 2.

In the final solution from Phase-1, node 3 is connected to nodes 4 and C (see Figure 2). If 3 is added into the network before 4 then it results in very high network cost (Equation 8) since the degree of node 3 will be less than 2. Hence the optimal solution should avoid adding node 3 before node 4. This eliminates 3 out of the 6 possible solutions (see Table 2). Close observation of the remaining 3 solutions from the dynamic programming output indicates the following: adding node 2 before node 4 results in very small degree and link utilization costs, but results in high SD time and TD costs due to high network convergence time after link failures. This could be attributed to the distribution of exit points in the network for various prefixes. Note that given different operational conditions (like traffic matrix and BGP prefix distribution), this condition may not be true. Finally adding node 3 before node 2 also resulted in high SD time and TD costs. Hence, for this example, the optimal solution is as shown in Table 3. However, ignoring SD time and TD costs in the DP formulation resulted in a different sequence for node addition. The cost for this sequence is much higher than the original sequence (see Table 2, Solution-5 and Solution-6). Note that although the sequence was determined by ignoring SD time and TD cost, the final cost includes the SD time and TD costs to emphasize the importance of considering user-centric performance metrics during the upgrade process.

It is important to note that the outcome discussed above depends on the operational network conditions and will vary depending on the state of the network. Our main focus in this paper is to show that the two-phase approach provides a framework that helps in finding an optimal strategy for network upgrade.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a two-phase framework to determine the optimal network upgrade strategy. To the best of our knowledge this is the first piece of work to address this problem. We showed that it is important to consider the performance perceived by customers during the upgrade process. We also used an example to show the feasibility and advantages of our proposed framework.

We have made several simplifying assumptions, but we believe that this is the first step towards developing a more comprehensive framework for the network upgrade process. In the future we plan to relax these assumptions. For example, the estimated traffic and revenue could dynamically change over time and could have several interdependencies. We also ignored the situation where the budget becomes available sequentially. Such changes can be easily incorporated into our framework by using adaptive and stochastic dynamic programming techniques that can adjust the optimal path after every stage based on the current state and the estimated values of other inputs. A detailed sensitivity analysis of our model due to variations in different parameters is also a part of our future work.

Another important point to note is that our framework does not require closed form expressions for network performance metrics (i.e. different constraints and costs). The same framework will still apply when the performance metrics need to be computed using complex algorithms. In practice, simulation techniques can be combined with our multistage dynamic programming framework to efficiently compute optimal solutions. We are investigating this further as a part of our ongoing research. We believe that the two-phase framework presented in this paper is powerful yet flexible. In the future we also plan to apply this framework to different networks such as wireless cellular networks.

6. REFERENCES

[1] Sprint ATL Tutorial. http://ipmon.sprint.com/pubs_trs/tutorials/Backbone-Architecture-And-Performance.pdf.

[2] J. Altmann and L. Rhodes. Dynamic Netvalue Analyzer: A Price Plan Modeling Tool for ISPs Using Actual Network Usage Data. In *IEEE Workshop on Advance Issues in E-Commerce and Web-Based Information Systems*, 2002.

[3] K. Curran, P. Flynn, and T. Lunney. A Decision Support System for Telecommunications. *International Journal of Network Management*, (12), 2002.

[4] P. Forsyth, Y. dHalluin, and K. R. Vetzal. Managing Capacity for Telecommunications Networks under Uncertainty. *IEEE/ACM Transactions on Networking*, 10(4), Aug. 2002.

[5] G. Iannaccone, C. Chuah, R. Mortier, S. Bhattacharyya, and C. Diot. Analysis of Link Failures in an IP Backbone. In *ACM IMW*, Nov. 2002.

[6] S. Jagannathan, J. Altmann, and L. Rhodes. A Revenue-based Model for Making Resource Investment Decisions in IP Networks. In *IEEE/IFIP Symposium on Integrated Network and System Management*, 2003.

[7] R. Keralapura, A. Moerschell, C. Chuah, G. Iannaccone, and S. Bhattacharyya. A Case for Using Service Availability to Characterize IP Backbone Topologies. *IEEE/KCIS Journal of Communications and Networks*, June 2006.

[8] A. Markopoulou, G. Iannaccone, S. Bhattacharrya, C. N. Chuah, and C. Diot. Characterization of Failures in an IP Backbone Network. In *IEEE Infocom*, Mar. 2004.

[9] K. Stordahl and E. Murphy. Forecasting Long-term Demand for Services in the Residential Market. *IEEE Communications Magazine*, 33, Feb. 1995.

[10] J. Xiao and R. Boutaba. Customer-centric Network Upgrade Strategy: Maximizing Investment Benefits for Enhanced Service Quality. In *IEEE/IFIP NOMS*, 2004.