Dynamic Grooming and RWA in Translucent Optical Networks Using a Time-Slotted ILP

X. Wang and M. Brandt-Pearce
Charles L. Brown Dept. of Elect. and Comp. Engineering
University of Virginia,Charlottesville, VA 22904
xw2bx@virginia.edu, mb-p@virginia.edu

S. Subramaniam
Dept. of Electrical and Computer Engineering
George Washington University, Washington, DC 20052
suresh@gwu.edu

***Abstract*—Translucent fiber-optic networks are carefully planned to achieve high capacity utilization efficiency as required by society's ever-increasing traffic demand. Existing research treats the problem of resource placement largely as a static design problem, which is solved with linear programming (LP) to find the optimal solution. The dynamic operational problem (grooming, regeneration, routing, and wavelength assignment) is approached using heuristic methods with the goal of improving the overall network performance given an existing network infrastructure. Our work combines these two approaches and solves a real-time dynamic traffic scenario with integer linear programming (ILP), seeking to maximize the overall network throughput. The traffic is served in a time-slotted fashion so that the network throughput is optimized at each time slot given the existing network state. The solution is compared with results from existing heuristic methods. We incorporate physical impairment limitations into our network model, and consider several grooming options.**

I. Introduction

Although the spectrum efficiency of optical networks has been greatly increased by the introduction of dense wavelength division multiplexing (DWDM), it is still of great research interest because of the dramatically-increased need for traffic volume from new applications. Translucent fiber-optic networks need to be carefully planned at many different stages so the capital investment can be minimized during network design and network upgrade, and the service capability can be maximized once the equipment is deployed and real-time traffic demands arrive.

One of the most powerful tools used to solve network planning problems is linear programming (LP), usually in the form an integer LP (ILP) or a mixed integer linear program (MILP). This approach has been applied to solve many of optical networking's key issues, such as: minimizing capital investment [1], [2], the regenerator placement problem [3], [4], network restoration [5], and network service capability maximization [6]. However, in these and other previous work when researchers use LP to address network problem they only consider a static traffic model. The dynamically fluctuating nature of real-time traffic cannot be properly represented with this kind of model. An effective alternative to using an LP method is to employ a heuristic method, which can handle dynamic environments. Heuristic methods usually have the advantage of a simple implementation, short calculation time, and good scalability. The drawback of heuristic methods is that they often do not provide optimal solutions. This is partly due to the fact they do not consider all possible solutions, and often pick one based on the network designer's understanding of the current network state.

In order to solve the network operation problem for dynamic real-time traffic and find a solution closer to the optimal solution, we propose to solve an ILP in a time-slotted fashion: for each time slot the network throughput is maximized while preserving as much of the resources as possible for future traffic. This requires a different objective function than previous LP work. The dynamic traffic model differs from the static traffic model in that each call request has a arrival time and a call duration, and the network capacity is updated continuously. Our approach bundles incoming traffic into time slots and solves the grooming, routing and wavelength assignment (GRWA) problem for those incoming calls optimally given the existing network state. We incorporate physical impairments (PIs) into our solution by imposing all-optical reach constraints. We also allow for multiple line-rates across the network. Another aspect that differentiates our approach from others is that we incorporate grooming into our solution in real time. This operation requires unique ILP constraints so that various grooming rules are enforced.

The rest of this paper is organized as follows. Section II discusses the regeneration model and the grooming model in our approach, both of which introduce constraints into our ILP model. Section III gives the detail of the ILP, and Section IV shows numerical result based on a network example. Section V concludes the article.

II. QoT, Regeneration and Grooming

As the optical signal travels through the network it suffers many physical impairments (PI) such as attenuation, noise, inter symbol interference (ISI), and nonlinear effects. The signal quality is thus degraded along the lightpath. Usually data transmission requires a specified bit error rate (BER) ($BER \leq 10^{-3}$ before error-control coding is a common standard). This is referred to as a quality-of-transmission (QoT) requirement. One way to compensate for these physical impairments is to implement signal regenerators along the lightpath. Signals undergo optical-electrical-optical (OEO) conversion, i.e., a clean pulse-shape is regenerated in the electrical domain. This operation is expensive since high-speed electronic devices are primary contributors to network cost. Therefore, regeneration sites need to be carefully selected and OEO circuits must be used sparingly so the network can operate with the lowest cost.

The distance that an optical signal can travel without regeneration while maintaining an acceptable QoT is called the transmission reach (TR). Since the PIs depend on the wavelength's line-rate, the TR also depends on the line-rate, even when other factors are constant. In this paper we use regenerator placement and TR results derived in [7]. In [7] four different regenerator placement algorithms are compared. The most powerful one is a heuristic algorithm called the 'signal quality prediction based (SQP) algorithm', which takes the TR as a constraint to search for regenerators allocation. Besides the regeneration function, regenerators also provide a wavelength conversion function otherwise unavailable to the lightpath. In our model, a wavelength continuity constraint only applies to the all-optical portion of a lightpath, i.e., between regeneration operations, the source, or the destination of the demand. We refer to the all-optical portion of a lightpath as a *segment*.

Grooming is an operation that takes low data-rate sub-streams and multiplexes them together onto the same wavelength to form a high data-rate stream. This operation can currently only be done in the electrical domain, which implies that we can only perform grooming either at the source of a demand or at an OEO regeneration site of a demand, which are the only two places where the signal reverts to its electrical form. This restriction requires us to carefully track the wavelength availability of channels for grooming opportunities. A wavelength must be considered as unavailable to groom onto along an all-optical segment, where it remains in the optical domain. To address this constraint, a novel approach must be applied. It is essential in our formulation to clearly mark all-optical segments that include more than one physical link so that OEO conversion is forbidden within the segment. Our approach is to create a virtual link between the two ends of a segment whenever it is formed, while marking the original physical resource unavailable to all new traffic demands. The virtual link holds only the wavelength used in the segment and its free capacity is adjusted based on the channel usage on the segment. New arrivals only see the one wavelength on these virtual links, and it will appear as a single hop to prevent a call from attempting to groom mid-way. When the routing and wavelength assignment ILP uses this virtual link, it considers only solutions that include utilizing that wavelength from the segment source to the segment end, but not partially in between. This ensures that grooming is only performed at points where signals undergo OEO conversion. This modification only lasts until the termination of the call that forms the segment.

III. Mathematical model

To the author's knowledge, this is the first published work that solves the grooming problem for real-time traffic using an ILP. Using an ILP allows us to fully exploit the versatility and routing potential of the network by considering all possible solutions, as compared to heuristic methods that consider only a small set of candidate paths. Our approach also differs from ILP approaches that also restrict the number of candidate paths for each node pair, such as [8].

A. Time-Slotted Approach

The novelty of our approach stems partly from viewing the dynamic problem as a succession of small static problems with strong initial conditions (the current network state). We solve the real-time dynamic traffic situation by using a time-slotted approach, applying the ILP to each time-slot successively. Similar to the dynamic heuristic method, the network state is updated continuously and forms a base for the ILP execution for next time slot.

Fig. 1. Illustration of time-slotted traffic model. Time slot duration of (a) 1 time unit and (b) 0.2 time units.

The pooling of traffic arriving within a time slot provides the algorithm flexibility to find an optimal GRWA solution for those calls, i.e., optimized for that time slot. In order to take the nature of the dynamic traffic into consideration, we split the time axis during which traffic arrives into fixed time slots, which creates a discrete-time approximate model of the real-time traffic, as shown in Fig. 1. In this approximation the demands is always serviced starting at the beginning of a time slot and terminates at the end of a time slot. This can be thought of as introducing a certain amount of delay (no longer than the duration of the time slot) for the demand to be serviced, called the initiation delay. The result is also an over-allocation of resources for demands after their termination (termination delay), again no longer than the duration of the time slot. By doing this, we create an opportunity to find an optimal solution for all the calls that arrive within the same

TABLE I
SETS USED BY ILP

N^{sd}	Set of network nodes
nbN	Number of nodes in the network
N^r	Set of nodes that are able to regenerate signal
N^{nr}	Set of nodes that are not capable of regeneration or those with the capability but not selected for regeneration
L	Set of network links
D	Set of network demands between node pairs. These demands often ask for data rate lower than line rate of a wavelength
Ld	Set of available wavelengths on each fiber link.

time slot. Fig. 1 illustrates the call initialization delay and call termination delay suffered by calls, which depends on the time slot duration. The time slot duration also affects the number of calls to be calculated per ILP solution. In this example, when the time slot duration is set to 1 time unit, Fig. 1(a), 5 calls arrive within 2 time slots (around 2.5 calls per time slot); when the time slot duration is set to 0.2 time units, Fig. 1(b), 5 calls arrive within 8 time slots, which makes it on average 0.625 calls per time slot. The delay introduced in servicing calls can be considered a trade-off for system performance improvement.

At the beginning of each time slot, the calls that have arrived in the previous time slot are used as demand requests and the solution is found based on the current network state (wavelength usage, etc). The network state is then updated based on the ILP's GRWA solution. The available capacity of each wavelength is updated. The ILP enforces the restriction of no wavelength conversion for all-optical segments. Then, before the next time slot starts, call terminations are performed, where some resources are released and the network state is again updated. The above steps are performed iteratively until all traffic samples terminate. The overall goodput (fraction of the data rate request accepted) is calculated by dividing the sum data-rate of traffic accepted by the system by the sum data-rate of the traffic requested.

B. ILP model

In this section we introduce the ILP model to optimize the network goodput by optimally aggregating sub-wavelength data traffic into high data-rate streams. The notation in Tables I, II, and III is used in our mathematical formulation. In particular we use the symbols i and j to index the head and tail of a physical link, and symbols s and d to index the source and destination of a demand.

Objective function:

$$\max \sum_{sd \in D} Accept_{sd} \times Datarate_{sd} + \sum_{ij \in L, ld \in Ld} ULLd_{ij,ld} \times w_{ld} * ld + \sum_{ij \in L, ld \in Ld, sd \in D} U_{ij,ld,sd} \times w \quad (1)$$

TABLE II
PARAMETERS USED BY ILP

$A_{ij,ld}$	how much of channel capacity is available of each link ij and ld pair, $ij \in L, ld \in Ld$
R_n	the number of regenerators allocated at node n
Tr_{ld}	Transmission reach for each wavelength, this is based on the channel line rate of that wavelength
$Datarate_{sd}$	Data rate that is requested by demand sd in D
$Length_{ij}$	Link length of link ij in kms
C_{ld}	Channel capacity (i.e.: channel line rate) of each wavelength
$S_{n,sd}$	$S_{n,sd} = -1$ if node n is demand sd's source node; $S_{n,sd} = 1$ if node n is demand sd's destination node; otherwise, $S_{n,sd} = 0$.
w	small negative weight parameter ≈ -0.01 to be put on each link-lambda pair usage; therefore, the shortest path routing is followed when there exist multiple equivalent solutions (equal throughput).
w_{ld}	small negative weight parameter ≈ -0.001 put on wavelength ld. Different values are used to enforce preferences for particular wavelengths; for example, a higher w_{ld} is assigned to wavelengths with low channel capacity to allocate the wavelength with the minimum sufficient capacity to a demand first.

TABLE III
VARIABLES USED BY ILP

$Accept_{sd}$	$Accept_{sd} = 1$ if demand sd is accepted; otherwise $Accept_{sd} = 0$
$U_{ij,ld,sd}$	$U_{ij,ld,sd} = 1$ if wavelength ld on link ij is assigned to demand sd
$UL_{ij,sd}$	$UL_{ij,sd} = 1$ if link ij is assigned to demand sd; otherwise $UL_{ij,sd} = 0$
$ULLd_{ij,ld}$	number of demands that share wavelength ld on link ij.
$\delta_{ij,sd,ld}$	physical distance from the head of link ij to the head of the segment for demand sd and wavelength ld. $\delta_{ij,sd,ld} = 0$ if wavelength ld on link ij is not assigned to demand sd.
$Y_{n,sd,ld}$	physical distance from node n to the head of the segment for demand sd and wavelength ld. $Y_{n,sd,ld} = 0$ if node n is not on the lightpath selected for demand sd.

The objective function is designed to maximize the demands serviced while using as few network resources as possible. The network resource usage is used as a tie breaker when multiple solutions result in the same goodput.

Constraints:
Conservation flow constraint:

$$\sum_{ij \in L, j=n} UL_{ij,sd} - \sum_{ij \in L, i=n} UL_{ij,sd} = S_{n,sd} \times Accept_{sd}, \quad \forall n \in N, sd \in D \quad (2)$$

Link capacity constraint:

$$\sum_{sd \in D} U_{ij,ld,sd} \times Datarate_{sd} \leq C_{ld} \times A_{ij,ld}, \forall ij \in L, ld \in Ld \quad (3)$$

No call splitting:

$$\sum_{ld \in Ld} U_{ij,ld,sd} \leq 1, \forall sd \in D, ij \in L \quad (4)$$

Wavelength continuity constraint:

$$\sum_{ij \in L, j \neq s, j \neq d} U_{ij,ld,sd} = \sum_{ij \in L, i \neq s, i \neq d} U_{ij,ld,sd}, \quad \forall sd \in D, ld \in Ld \tag{5}$$

Transmission reach constraints, $\forall sd \in D, ij \in L, ld \in Ld$:

$$\begin{aligned} \delta_{ij,sd,ld} &\leq U_{ij,ld,sd} \times Tr_{ld}, \\ \delta_{ij,sd,ld} &\leq Y_{i,sd,ld} \\ Y_{i,sd,ld} - \delta_{ij,sd,ld} &\leq Tr_{ld} \times (1 - U_{ij,ld,sd}) \\ Y_{n,sd,ld} &= 0, \forall n \in N^r \text{ or } n = s \\ Y_{n,sd,ld} &= \sum_{ij \in L: j=n} \delta_{ij,sd,ld} + Length_{ij} \times U_{ij,ld,sd}, \\ & \qquad \forall n \in N^{nr} \text{ and } n \neq s \end{aligned} \tag{6}$$

Regenerators resource constraint:

$$\sum_{sd \in D, nj \in L, ld \in Ld} U_{nj,ld,sd} \leq R_n, \forall n \in N^r \tag{7}$$

Link-wavelength usage count:

$$ULLD_{ij,ld} = \sum_{sd \in D} U_{ij,ld,sd}, \forall ij \in L, ld \in Ld \tag{8}$$

Link usage with link-wavelength usage:

$$UL_{ij,sd} = \sum_{ld \in Ld} U_{ij,ld,sd}, \forall ij \in L, sd \in D \tag{9}$$

No grooming constraint (optional):

$$\sum_{sd \in D} U_{ij,ld,sd} \leq 1, \forall ij \in L, ld \in Ld \tag{10}$$

C. Explanation of Constraints

Equation (2) ensures that for transient nodes (neither the source nor the destination for the demand considered) the traffic entering the node equals the traffic exiting the node. For end nodes, the traffic that exits the source node and enters into the destination node equals the demand's datarate request. Equation (3) ensures that, for the wavelength (lambda) considered, the traffic volume assigned to new demands is less than the available capacity. Note that a link-lambda pair could be used for one demand and later also assigned to another demand if there is capacity left over; so $A_{ij,ld}$ ranges from 0 to 1. Equation (4) ensures that each demand request can only be satisfied by at most one lambda on any link. This assumption is solely for the sake of model simplicity, and could be removed to consider an expanded version of the problem. Equation (5) ensures that for each segment the optical signal maintains the same lambda. The constraint is placed on the nodes that do not have regeneration capability (also on nodes with regeneration capability if the regeneration function is not used). For these nodes the traffic that enters and the traffic that exits the node for a demand use the same lambda.

Equations (6) ensure that the optical signal will not travel further than the transmission reach without regeneration: if a link is assigned to a demand, its head node should be within the transmission reach from the segment's head node. The distance from the segment head node to a link ij should be no greater than the distance to the node i. If a link is assigned to a demand ($U_{ij,ld,sd} = 1$), $Y_{i,sd,ld}$ (abbreviated Y) and $\delta_{ij,sd,ld}$ should be equal. If it is not assigned to the demand, Y should be within the transmission reach. For the source node of the demand or a node performing a regeneration operation for that demand, Y is 0. This is how the Y's are derived from the source of the demand. If a link is assigned to the demand, its end node Y is calculated by adding the physical link distance to the Y value of its head node. For any node along the lightpath for a demand, the Y should be within the transmission reach.

Equation (7) enforces that the total number of OEO regenerators at any given regeneration node at any given time slot is no greater than the number of regenerator circuit the node has. Equations (8) and (9) show the relation between the three variables we use in our programming. Equation (10), which enforces how many demands can share a wavelength ld on a link ij, is optional and is used to compare the grooming case versus the no-grooming case.

IV. Numerical Results

Numerical simulation experiments are conducted on a 14-node bidirectional-link NSF nationwide network, shown in Fig. 2. Each link has 8 wavelengths, and each wavelength supports either a 10, 40 or 100 Gbps line-rate and has unlimited, 2500 and 2000 kms TR, respectively. The ILP is solved using the optimizer CPLEX [9]. The demand requests arrive at the system following a Poisson process and are uniformly distributed among node pairs. The call duration follows an exponential distribution with a mean value of 1 (arbitrary time units). For Figs. 3-8, the time slot duration is 0.1 time units. The data-rate requests for each demand follow a uniform distribution ranging from 1 to 30 Gbps.

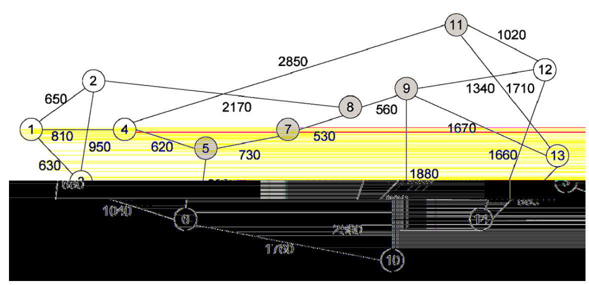

Fig. 2. 14-node NSF nationwide network

Since in real-time scenarios the processing time is critical, the ILP solve-time is limited to 100 seconds (wall-time) on a desktop computer; when the calculation time exceeds this limit, the best feasible solution at that time is selected.[1] The computation time is therefore significant when the call-durations are short compared to this delay. With optimized

[1]We tested with different simulation time constraints and noticed that while it is important to have a solve-time sufficiently long to find at least one feasible solution, finding the optimal solution does not improve the system performance significantly.

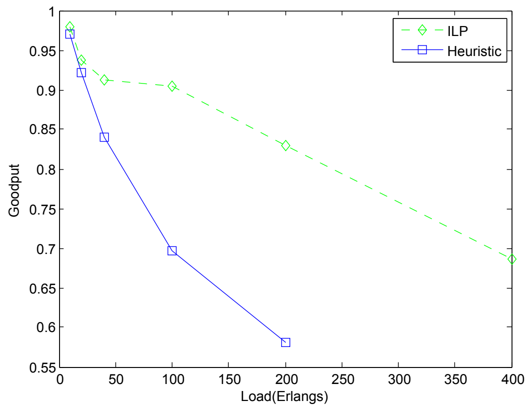

Fig. 3. Goodput using ILP and heuristic method versus load

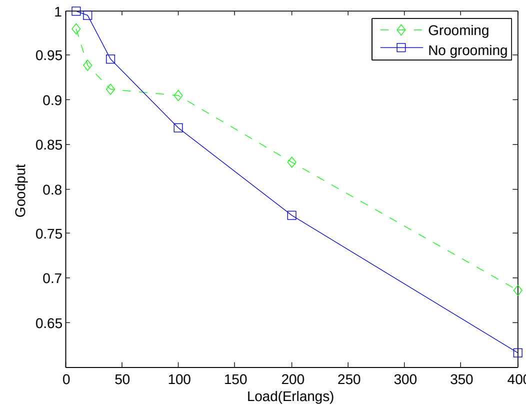

Fig. 5. Goodput using grooming and no-grooming methods versus load

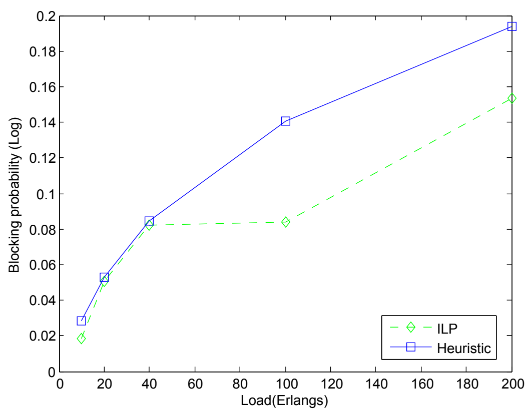

Fig. 4. Blocking probability using ILP and heuristic method versus load

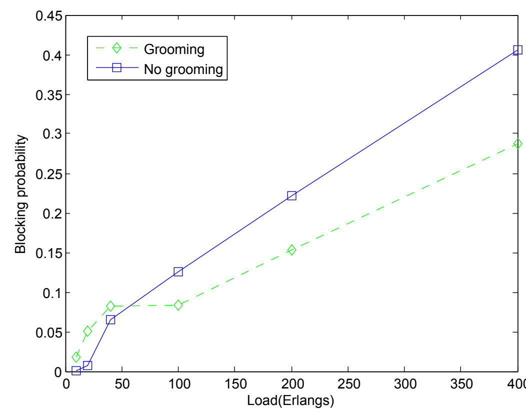

Fig. 6. Blocking probability using grooming and no-grooming methods versus load

processing the delay may be reduced significantly, yet the algorithm remains complex for large networks. The proposed technique, as stated, is suitable to calls with long service time requirements, such as large file transfers and other e-science applications.

We compare the solution from solving the ILP with the solution from a previously proposed heuristic method [10]. For a high load case, we can see significant improvements in terms of both network goodput, Fig. 3, and call blocking probability, Fig. 4. This is because the ILP considers all possible lightpaths given by the network structure while the heuristic method often considers the shortest path or just a few alternatives before it rejects the request. It is obvious that the network operator benefits from the more comprehensive search within the solution space. From the plots we can see that for low loads the ILP method provides little advantage since the low traffic rarely causes congestion.

The grooming case outperforms the no-grooming case as the load increase, as seen in Figs. 5 and 6. In the no-grooming case, a channel is set as unavailable once assigned to a demand, no matter how much capacity is still available. If the call cannot groom with others, there is also a smaller chance for it to find a feasible lightpath when the network is heavily loaded. That is why we can see the grooming case outperforms the no-grooming case in our ILP solution.

For some cases (for low load) the performance improvement is not obvious or sometimes reversed. This is because the time-slotted ILP solution is based on an objective function to maximize the single time slot throughput, rather than the whole simulation throughput. In the future, we will consider improvement to address this issue.

In Figs. 7 and 8, we compare the difference when re-generation is allowed versus not allowed, i.e., between a transparent and a translucent network. If no regeneration is possible, the physical impairments put a restriction on the signal transmission reach, which then makes certain node pairs unreachable for high line-rates. The regeneration operation basically extends the transmission reach for each line-rate lightpath, thereby introducing more candidate solutions. For node pairs that are too far apart in terms of physical distance, it is possible to create a lightpath for some line-rates with the signal being regenerated somewhere along the lightpath. For heavily loaded networks, the direct connection (the shortest path) between a node pair may be unavailable, and the demand may be required to be rerouted over a longer lightpath, where the signal quality degradation could then become an issue. Regeneration makes rerouting possible.

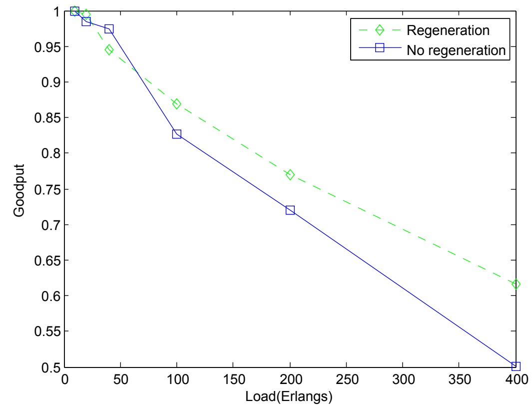

Fig. 7. Goodput using regeneration and no-regeneration methods versus load

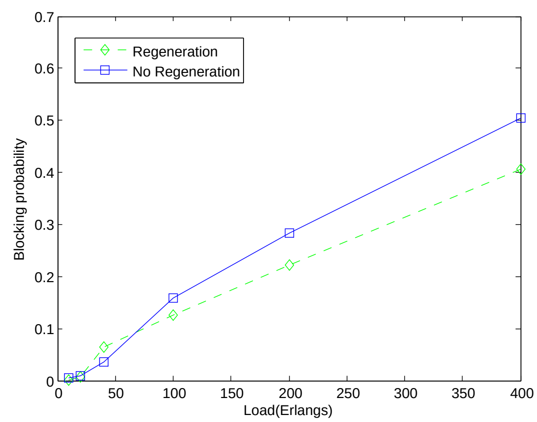

Fig. 8. Blocking probability using regeneration and no-regeneration methods versus load

TABLE IV
EFFECTS OF TIME SLOT DURATION, LOAD OF 100 ERLANG

Time slot duration	0.2	0.1	0.05	0.03
Goodput	0.92	0.91	0.88	0.91
Blocking probability	0.09	0.08	0.05	0.09

Another important impact on performance caused by regeneration is that it provides opportunities for grooming. When a signal undergoes OEO conversion, it can then be multiplexed with other streams to become a higher data-rate stream, which improves the channel capacity usage efficiency. That is why researchers propose to intentionally control TR of some line-rate signals to force them through OEO conversions, to take advantage of these grooming opportunities.

Note that although obvious differences in performance exist between different methods, the difference is not large. Since the ILP considers all options, the network is highly efficient and can handle large volumes of traffic.

In Table IV we compare the system performance for a traffic model with load of 100 Erlang using different time slot durations. The value of the time slot duration can be considered as the maximum delay of each call before it can be processed. From the ILP's point of view, it affects the number of simultaneous demands for service. On average, the number of calls per time slot is equal to the load (Erlang) times the slot duration divided by the call duration (arbitrary units of time). As the slot size increases, the number of calls to allocate per time slot increases. The ILP solution calculated with more calls results in a better solution, since the optimization only considers new calls (not calls established in previous slots). However, the increased time slot duration wastes resources since the termination of calls is delayed until the end of each time slot. Based on these two reasons, we can explain why the performance is similar with different time slot durations, as seen in the table.

V. CONCLUSION

We propose an approach to solve the dynamic real-time traffic GRWA problem using an ILP in a time-slotted fashion. The traffic only endures a short delay before being processed. Meanwhile, the system efficiency is increased compared to heuristic methods, both in goodput and call blocking probability. Grooming and regeneration effects on the network performance are also discussed.

Using an ILP in a dynamic system may become too computationally burdensome to be scalable as the network size increases. In future work we will consider relaxation techniques for obtaining faster near-optimal solutions using this approach.

VI. ACKNOWLEDGEMENT

This work was supported in part by NSF grants CNS-095795 and CNS-0916890.

REFERENCES

[1] M. Batayneh, D. A. Schupke, M. Hoffmann, A. Kirstaedter, and B. Mukherjee, “On routing and transmission-range determination of multi-bit-rate signals over mixed-line-rate WDM optical networks for carrier ethernet,” *IEEE/ACM Transactions on Networking*, 2011.

[2] A. Jarray and B. Jaumard, “Exact ILP solution for the grooming problem in WDM ring networks,” in *IEEE International Conference on Communications (ICC)*, May 2005.

[3] S. Bandyopadhyay, Q. Rahman, S. Banerjee, S. Murthy, and A. Sen, “Dynamic Lightpath Allocation in Translucent WDM Optical Networks,” in *IEEE International Conference on Communications (ICC)*, June 2009.

[4] Y. Chen and A. Jaekel, “Regenerator allocation for scheduled lightpath demands in translucent optical networks,” in *International Conference on Computing, Networking and Communications (ICNC)*, 2012.

[5] M. Jin, R.Q. Hu, Z. Zhang, and W. Hu, “Integer linear programming models and performance evaluation on wavelength rearrangement in a mesh-restored all-optical network,” *IEEE Communications Letters*, Feb. 2006

[6] G. Shen and R.S. Tucker, “Sparse Traffic Grooming in Translucent Optical Networks,” *IEEE/OSA Journal of Lightwave Technology*, Oct. 2009

[7] X. Yang and B. Ramamurthy, “Sparse regeneration in translucent wavelength-routed optical networks: architecture, network design and wavelength routing,” *Photonic Network Communications*, 2005.

[8] A.I. Al-Fuqaha, G.M. Chaudhry, M. Guizani and G.B. Brahim, “Routing in all-optical DWDM networks with sparse wavelength conversion capabilities,” in *IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2003

[9] IBM ILOG CPLEX Studio Academic Research 12.3

[10] X. Wang, M. Brandt-Pearce, and S. Subramaniam “Grooming and RWA in Translucent Dynamic Mixed-Line-Rate WDM Networks With Impairments” in *OFC conference*, March 2012