Solving Very Large RWA Data Instances

Brigitte Jaumard and Maryam Daryalal

Department of Computer Science and Software Engineering

Concordia University

Montreal (QC) H3G1M8

Email: bjaumard@cse.concordia.ca

Abstract—While the Routing and Wavelength Assignment (RWA) has been widely studied, few studies attempt to solve realistic size instances, namely, with 150 wavelengths per fiber and a few hundred nodes. Indeed, state of the art is closer to around 20 nodes and 30 wavelengths, regardless of what the authors consider, heuristics or exact methods with very few exceptions. In this paper, we are interested in reducing the gap between realistic data sets and testbed instances that are often used, using exact methods. Even if exact methods may fail to solve in reasonable time very large instances, they however output bounds on the accuracy of the heuristic solutions. The novelty of our work is to exploit the observations that optimal solutions contain a very large number of lightpaths associated with shortest paths or k shortest paths with a small k. We propose different RWA algorithms that lead to solve exactly or near exactly much larger instances than in the literature with up to 150 wavelengths and 50 nodes.

I. Introduction

Offline Routing and Wavelength Assignment (RWA) is the key provisioning problem of optical networks. The continuous growth of traffic, nourished by the emerging rich-content high-rate and bursty applications, such as video on demand, on-line gaming, high-definition television (HDTV), and cloud computing, can only be met with the abundant capacity provided by optical transport networks. The following numbers provide an idea of what growth of traffic means in 2016. For instance, Global Internet Protocol (IP) traffic has increased more than fivefold in the past 5 years, and will increase nearly threefold over the next 5 years according to a Cisco study [1]. A recent study (Durairajan *et al.* [2]) of the US long-haul fiber-optic infrastructure, based on the detailed fiber deployment maps from 5 tier-1 and 4 major cable providers (AT&T, Comcast, Cogent, EarthLink, Integra, Level3, Suddenlink, Verizon and Zayo), suggests that the US long-haul fiber network has 273 nodes/cities, 2411 links, and 542 conduits.

While the RWA problem was the focus of many studies in the 80's, numerous heuristics have been proposed. However, they have not been tested and compared on very large data instances except for very few papers, see, e.g., [3]–[6], and most of the time, they come without any information on the accuracy of their solutions. On the other hand, exact methods have difficulties to scale. The objective of the present paper is to revisit the ILP model based on independent sets [7], [8] and to increase its scalability. This can be done through the maximization of lightpaths with a route using a shortest path selected among all shortest paths for every node pair, rather than using a single shortest path selected at random, as in

many previous studies. Following the observation that even if the traffic among the node pairs are evenly distributed in the network, the link utilization varies from one link to the next, and the smaller is the number of idle wavelengths, the most likely the link is a bottleneck one, even if the link load does not fully provide information with respect to the wavelength continuity constraint, i.e, finding a lightpath so that we can use the same wavelength all the way towards the destination. In order to go around bottleneck links, we propose to use longer paths, among the k-shortest paths that are not shortest paths and that are not going through those bottleneck links.

Given a set of lightpath requests, two variants of the RWA problem have been studied in the literature: max-RWA and min-RWA. For the former problem, the objective is to maximize the number of lightpath requests that can be routed with a given number of wavelengths. For the latter one, the objective is to minimize the number of wavelengths to route all the requests.

Several surveys have been written on the RWA problem, where the reader can find a comprehensive survey of the various mathematical models that have been investigated, see, e.g., [9] for symmetrical traffic and [10] for asymmetrical traffic. Among the papers that explored exact solutions, we find three classes of Integer Linear Programming (ILP) formulations for the RWA problem: *(i)* link-based, *(ii)* path-based, or *(iii)* maximal independent set (MIS)-based. Comparison of link and path based can be found in [8], [9] with the objective of maximizing the grade of service.

The largest instances that have been solved exactly so far for mesh networks are the EON network (20 nodes, and 39 optical links) with 24 wavelengths, the Brazil network (27 nodes and 70 optical links) with 14 wavelengths, both for the GoS maximization [9]. In terms of heuristics, large instances have been solved by Martins *et al.* [5]: 26 realistic instances with up to either 90 nodes or 175 links, with no information on the distance of their solutions to an optimal one, but with comparisons with other experimental results.

The paper is organized as follows. In the next section, we first recall the decomposition model that is based on maximal independent sets, and then propose two alternate reformulations of the pricing problem that is in charge of generating augmented wavelength configurations, i.e., subsets of lightpaths routed on a given wavelength, which can contribute to the improvement of the incumbent RWA solution. Section IV discusses the proposed solution schemes that take

advantage of the different formulations of the pricing problems of the previous section. Numerical results are presented in Section V. Conclusions are drawn in the last section.

II. STATEMENT OF THE RWA PROBLEM

Consider a WDM (Wavelength-Division Multiplexing) optical network represented by a multigraph $G = (V, L)$ with node set V indexed by v, where each node is associated with a node of the physical network, and with link set L indexed by ℓ where each link is associated with a fiber link of the physical network: the number of links from v to v' is equal to the number of fibers supporting traffic from v to v'. Connections and fiber links are assumed to be directional, and the traffic to be asymmetrical. The set of available wavelengths is denoted by Λ, and is indexed by λ with $W = |\Lambda|$. The traffic is defined by a $n \times n$ matrix D where D_{sd} defines the number of requested connections (i.e., wavelengths) from v_s to v_d. All wavelengths are assumed to have the same transport capacity. Let $\mathcal{SD} = \{(v_s, v_d) \in V \times V : D_{sd} > 0\}$. Let $\omega^+(v)$ (resp. $\omega^-(v)$) be the set of outgoing (resp. incoming) fiber links at node v. We assume that the same wavelength is used from the source to the destination for all connection requests. Note that it has been shown (see [11], [12]) that wavelength conversion (i.e., multiple-hop connections) does not help very much in order to reduce the blocking rate (max-RWA problem).

The RWA problem can then be formally stated as follows: given a multigraph G corresponding to a WDM optical network, and a set of requested connections, find a suitable lightpath (p, λ) for each granted connection, where a lightpath is defined by the combination of a routing path p and a wavelength λ, so that no two paths sharing a link of G are assigned the same wavelength. We study the objective of minimizing the blocking rate, that is equivalent to maximizing the number of granted connections (also called Grade of Service or GoS for short), leading to the so-called max-RWA problem.

III. A DECOMPOSITION OPTIMIZATION MODEL

Several authors have already investigated modelling the RWA problem with a decomposition model, within the framework of exact solution schemes, see, e.g., [7], [8]. We revisit those models here, with the goal of enhancing them in order to solve much larger RWA instances. We first recall the decomposition optimization model based on maximal independent set, as initially proposed by [7] and improved by [8], and then discuss how to improve it further.

A. Wavelength Configurations

Lee *et al.* [7] as well as Jaumard *et al.* [8] introduced the concept of independent routing configurations, where each configuration is associated with a set of paths that can be used for satisfying a given fraction of the connections with the same wavelength. Within a wavelength configuration, routes must be independent, in the sense that they must not share a link as two lightpaths can not share a link if they use the same wavelength.

An independent routing configuration c can be formally represented by a non negative vector a^c such that $a^c_{sd} =$ number of connection requests from v_s to v_d that are supported by configuration c, with $a^c_{sd} \leq D_{sd}$ for $(v_s, v_d) \in \mathcal{SD}$.

We denote by C the set of all possible wavelength configurations.

B. Decomposition RWA Model

We use two sets of variables. The first set of variables, z_c, enable the selection of the best configurations and of their number of occurrences (i.e., to how many wavelengths they apply). The second set of variables, y_{sd}, compute the GoS for each node pair (v_s, v_d), so that their sum provide the overall GoS.

The basic model is written as follows. Assuming that the configurations are at hand, the model selects the best wavelength configurations to maximize the grade of service (GoS), i.e., the number of granted connections.

$$\max \sum_{(v_s, v_d) \in \mathcal{SD}} y_{sd} \qquad (1)$$

$$\text{subject to: } \sum_{c \in C} z_c \leq W \qquad (2)$$

$$y_{sd} \leq \sum_{c \in C} a^c_{sd} z_c \qquad (v_s, v_d) \in \mathcal{SD} \qquad (3)$$

$$y_{sd} \leq D_{sd} \qquad (v_s, v_d) \in \mathcal{SD} \qquad (4)$$

$$z_c \in \mathbb{Z}^+ \qquad c \in C \qquad (5)$$

$$y_{sd} \geq 0 \qquad (v, v_d) \in \mathcal{SD}. \qquad (6)$$

Observe that, since $D_{sd} \in \mathbb{Z}^+$, we have $y_{sd} \in \mathbb{Z}^+$. However, we do not need to explicitly enforce it.

Constraints (3) ensure that we do not select more wavelength configurations than the number of available wavelengths. Constraints (4) prevent from granting more connections than requested. Constraints (6) and (5) define the domains of the variables.

IV. SOLUTION PROCESS AND ALGORITHMS

A. Implicit Enumeration of the Wavelength Configurations

The model proposed in Section III-B has an exponential number of variables, and therefore is not scalable if solved using classical ILP (Integer Linear Programming) tools. Indeed, we need to use column generation techniques in order to manage a solution process that only requires an *implicit* enumeration of the wavelength configurations (interested readers may refer to Chvatal [13]). Column generation method allows the exact solution of the linear relaxation of model (1) - (6), i.e., where constraints $z_c \in \mathbb{Z}^+$ are replaced by $z_c \geq 0$, for $c \in C$. It consists in solving alternatively a restricted master problem (the model of Section III-B with a very limited number of columns/variables) and the pricing problem (generation of a new wavelength configuration) until the optimality condition is satisfied (i.e., no wavelength configuration with a negative reduced cost). In other words, when a new wavelength

configuration is generated, it is added to the current restricted master problem only if its addition implies an improvement of the optimal value of the current restricted master problem. This condition, indeed an optimality condition, can be easily checked with the sign of the reduced cost, denoted by $\overline{\text{COST}}$, see (7) for its expression, of variables z_c.

Once the optimal solution of the LP (Linear Programming) relaxation ($z_{\text{LP}}^\star$) has been reached, we solve exactly the last restricted master problem, i.e., the restricted master problem of the last iteration in the column generation solution process, using a branch-and-bound method, leading then to an ε-optimal ILP solution ($\tilde{z}_{\text{ILP}}$), where

$$\varepsilon = \frac{\tilde{z}_{\text{ILP}} - z_{\text{LP}}^\star}{z_{\text{LP}}^\star}.$$

Branch-and-price methods can be used in order to find optimal solutions, if the accuracy (ε) is not satisfactory, see, e.g., [8], [14].

B. ε-Optimal Algorithms CG, CG+ and CG++

In the context of the present study, we will investigate different algorithms, in which the differences lie in the generation process of new augmenting configurations, i.e., of configurations that give rise to an improvement of the value of the current restricted master, when solving its linear relaxation.
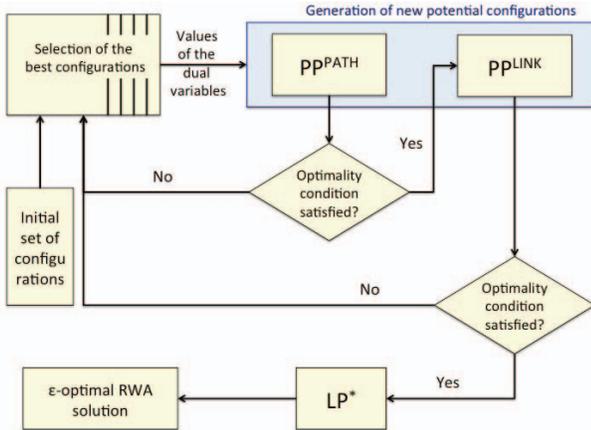

Fig. 1. Generic flowchart for all algorithms

The first algorithm, denoted by CG, will use a link formulation, for the mathematical model involved in the generation of new augmenting wavelength configurations. In the context of column generation, the configuration generator corresponds to the so-called pricing problem. The pricing problem with a link formulation will be called PP$_{\text{LINK}}$ (see Section IV-C for its detailed description), and corresponds to the algorithm used in [8]. The flowchart of algorithm CG is represented in Figure 1, under the assumption that box PP$_{\text{PATH}}$ is omitted.

Based on the observation made by several researchers, and investigated later in Section V, that a very high percentage of lightpaths are supported by shortest paths, or k-shortest paths [15] with a small k, we propose to investigate a path

formulation, called PP$_{\text{PATH}}$, for the pricing problem (see Section IV-D for its detailed description), with different strategies for selecting the paths. Since we cannot consider all possible paths, otherwise the pricing problem would not be scalable, we need to combine the use of PP$_{\text{PATH}}$ with PP$_{\text{LINK}}$ in order to get an ε-optimal algorithm. Indeed, when PP$_{\text{PATH}}$ is no more able to output an augmenting wavelength configuration, we switch to PP$_{\text{LINK}}$, and check whether it is still possible to generate an augmented wavelength configuration using more diverse paths than those considered in PP$_{\text{PATH}}$. Flowchart of the corresponding algorithm is represented in Figure 1.

We investigated two variants for PP$_{\text{PATH}}$. In the first one, we consider only the shortest paths, and in the second one, we consider the shortest paths, as well as a subset of at most 15 second shortest paths. By second shortest paths, we mean paths with a value larger than the one of the shortest paths, but limited to the first largest value. For instance, consider a pair of nodes such that the 10-shortest paths have values 7, 7, 7, 9, 9, 9, 9, 9, 12, 14. In such a case, we will consider all the shortest paths, i.e., the first 3 paths, and then the 5 second shortest paths, all with value 9. The resulting ε-optimal algorithms are called CG+ and CG++, respectively.

C. Pricing Problem - Link Formulation

As always with the column generation method, the objective of the pricing problem (i.e., generator of new wavelength configurations) is the reduced cost ($\overline{\text{COST}}_c^{\text{LINK}}$) of variable z_c. In order to alleviate the notations, index c will be omitted in the remainder of this section.

Let $u^{(2)}$ and $u_{sd}^{(3)}$ be the values of the dual variables associated with constraints (2) and (3) in the optimal solution of the linear relaxation of the current restricted master problem (see the flowchart in Figure 1). Consider the following set of variables:
$\alpha_\ell^{sd} = 1$ if link ℓ is used in a route from v_s to v_d, 0 otherwise. The link formulation of the pricing problem can be written as follows:

Wavelength Configuration Generator - Link Model PP$_{\text{LINK}}$

$$\max \quad \overline{\text{COST}}^{\text{LINK}} = -u^{(2)} + \sum_{(v_s,v_d)\in\mathcal{SD}} \sum_{\ell\in\omega^+(v_s)} \alpha_\ell^{sd} u_{sd}^{(3)} \quad (7)$$

subject to:

$$\sum_{(v_s,v_d)\in\mathcal{SD}} \alpha_\ell^{sd} \leq 1 \qquad\qquad \ell \in L \qquad\qquad (8)$$

$$\sum_{\ell\in\omega^+(v)} \alpha_\ell^{sd} = \sum_{\ell\in\omega^-(v)} \alpha_\ell^{sd} \qquad (v_s,v_d) \in \mathcal{SD},$$

$$v \in V \setminus \{v_s, v_d\} \quad (9)$$

$$\sum_{\ell\in\omega^+(v_s)} \alpha_\ell^{sd} \leq D_{sd} \qquad (v_s,v_d) \in \mathcal{SD} \quad (10)$$

$$\sum_{\ell\in\omega^-(v_s)} \alpha_\ell^{sd} = \sum_{\ell\in\omega^+(v_d)} \alpha_\ell^{sd} = 0 \qquad (v_s,v_d) \in \mathcal{SD} \quad (11)$$

$$\alpha_\ell^{sd} \in \{0,1\} \qquad\qquad \ell \in L, (v_s,v_d) \in \mathcal{SD}. \quad (12)$$

Constraints (8) ensure wavelength continuity, i.e., that a link cannot be traversed by more than one route in any given wavelength configuration. Routes are established with the help of the flow conservation constraints (9): if no route is selected for node pair (v_s, v_d), then $\alpha_\ell^{sd} = 0$ for all links $\ell \in L$, otherwise, the sum of the outgoing flow values at the source node ($\sum_{\ell \in \omega^+(v_s)} \alpha_\ell^{sd}$) gives the number of link-disjoint routes from v_s to v_d in the wavelength configuration under construction. Constraints (11) prevent loops around the source or the destination nodes from arising. Constraints (12) define the domain of variables α_ℓ^{sd}.

Correspondence between variables of the pricing problem and coefficients of the master problem:

$$a_{sd} = \sum_{\ell \in \omega^+(v_s)} \alpha_\ell^{sd}. \qquad (13)$$

Observe that nothing prevents from selecting several pairwise link disjoint paths for a given pair (v_s, v_d) of source and destination nodes. Indeed, a_{sd} is equal to the number of link-disjoint paths from v_s to v_d in the configuration under construction.

D. Pricing Problem - Path Formulation

In the path formulation, we provide a set P_{sd} of paths for each source and destination pair of nodes. We denote by P_{sd}^k the set of paths with the k smallest value (assuming we only care for distinct values). Consequently, P_{sd}^1 is the set of shortest paths for node pair (v_s, v_d), and P_{sd}^2 is the set of paths with the second distinct smallest value for the paths from v_s to v_d, and so on.

The path formulation for the wavelength configuration generator is denoted by PP$_{\text{PATH}}$. It uses the set of decision variables:
$\beta_p^{sd} = 1$ if path p is used in the wavelength configuration under construction, 0 otherwise.
PP$_{\text{PATH}}$ is written as follows:

Wavelength Configuration Generator - Path Model PP$_{\text{PATH}}$

$$\max \quad \overline{\text{COST}}^{\text{PATH}} = -u^{(2)} + \sum_{(v_s, v_d) \in \mathcal{SD}} \sum_{p \in P_{sd}} \beta_p^{sd} u_{sd}^{(3)} \quad (14)$$

subject to:

$$\sum_{(v, v_d) \in \mathcal{SD}} \sum_{p \in P_{sd}} \delta_\ell^p \beta_p^{sd} \leq 1 \qquad \ell \in L \qquad (15)$$

$$\sum_{p \in P_{sd}} \beta_p^{sd} \leq D_{sd} \qquad (v_s, v_d) \in \mathcal{SD} \qquad (16)$$

$$\beta_p^{sd} \in \{0, 1\} \qquad \ell \in L, (v_s, v_d) \in \mathcal{SD}. \quad (17)$$

Pairwise link disjointness for paths is guaranteed thanks to constraints (15), in which δ_ℓ^p is a binary value representing the presence of link ℓ in path p. Constraints (16) enforce not to exceed the lightpath demand. Constraints (17) define the domain of variables α_ℓ^{sd}.

TABLE I
CHARACTERISTICS OF THE DEMANDS

NSFNET				USANET			
Data	$|\mathcal{SD}|$	D	W	Data	$|\mathcal{SD}|$	D	W
NSF_0	141	436	30	USA_0	455	1,336	75
NSF_1	157	536	35	USA_1	508	1,611	85
NSF_2	173	632	35	USA_2	526	1,887	100
NSF_3	178	737	45	USA_3	534	2,161	115
NSF_4	180	827	45	USA_4	541	2,422	125
NSF_5		921	50	USA_5	546	2,706	140
NSF_6		1,015	55	USA_6	551	2,972	
NSF_7	181	1,105	60	USA_7		3,242	
NSF_8		1,198	65	USA_8		3,509	150
NSF_9		1,280	70	USA_9	552	3,803	
NSF_10		1,373	75	USA_10		4,086	
NSF_11		1,475	80				
NSF_12		1,571	85	GERMANY			
NSF_13		1,662	90				
NSF_14	182	1,752	95	GER_0		2,365	100
NSF_15		1,845	100	GER_1		3,041	130
NSF_16		1,933	100	GER_2	660	3,668	
NSF_17		2,027	105	GER_3		4,320	
NSF_18		2,100	110	GER_4		4,989	150
NSF_19		2,194	115	GER_5		5,666	

Correspondence between variables of the pricing problem and coefficients of the master problem:

$$a_{sd} = \sum_{p \in P_{sd}} \beta_p^{sd}.$$

V. NUMERICAL RESULTS

We now present the numerical experiments that were conducted in order to validate and test the performance of the proposed ε-optimal algorithms: CG, CG+, and CG++. In Section V-B, we present intensive numerical results, comparing the performance of all 5 algorithms, on three data sets.

A. Data Sets

We run experiments on three different topologies: the 14-node, 42-(directed) link NSFNET [16], the 24-node, 88-(directed) link USANET [17] and the 50-node, 176-(directed) link GERMANY [16].

For each network topology, we consider various traffic instances. For the first traffic instances (i.e., NSF_0 or USA_0) of the first two topologies, the directed traffic demand matrix $T = [T_{sd}]$ is generated by drawing the (integer) traffic demands (in units of lightpaths) uniformly at random in $\{0, 1, 2, 3, 4, 5\}$. For the GERMANY topology, the first traffic instance (i.e., GER_0) is from the snd·lib library [16]. The following traffic instances correspond to incremental traffic: NSF/USA/GER$_i$ ⊆ NSF/USA/GER$_{i+1}$ where NSF/USA/GER$_{i+1}$ is built upon NSF/USA/GER$_i$ by randomly adding ALEA more requests for each pair of nodes, in which ALEA is drawn uniformly from $\{1, 2, 3, 4, 5\}$ for NSF/USA and from $\{0, 1, 2, 3\}$ for GERMANY.

Table I gives the detailed characteristics of the request sets. For each traffic instance, we provide the number of node pairs with requests ($|\mathcal{SD}|$) and the overall number of traffic requests ($D = \sum_{\{v_s, v_d\} \in \mathcal{SD}} d_{sd}$). Note that 182, 552 and 2,450 are the maximum numbers of (ordered) node pairs, for NSFNET, USANET, and GERMANY, respectively.

B. Comparative Results

1) Parameters for the Algorithm Comparisons: In the next six tables, Tables II to VII, we report the results on an

increasing number of requests, and accordingly, an increasing number of wavelengths so that the GoS remains above (or close enough to) 80%. We first provide a lower bound ($z^{\star}_{\text{LP}}$, the optimal value of the linear relaxation of (2) - (6)) and an upper bound ($\tilde{z}_{\text{ILP}}$, the value of the ILP solution) on the optimal grade of service, so that the comparison of the two bounds leads to an estimate of the accuracy of the GoS obtained with the different algorithms. In addition, we provide #PP$_{\text{LINK}}$, i.e., the number of generated wavelength configurations before reaching the optimality condition in the solution of the linear relaxation of model (2) - (6). Lastly, for each algorithm, we give the computational times, expressed in seconds.

For CG+ and CG++ algorithms, the second and third ε-optimal algorithms, we provide #PP_iter as the sum of two numbers, #PP_iter (PATH), the number of wavelength configurations generated by PP$_{\text{PATH}}$, and #PP_iter (LINK), the number of wavelength configurations generated by PP$_{\text{LINK}}$. In addition, for CG++, we provide the overall number of paths from $\bigcup_{(v_s,v_d)\in\mathcal{SD}} (P^1_{sd} \cup P^2_{sd})$, which are used in the selected wavelength configurations.

CG except for the NSF_15 data instance. Improved values account for between 1 and 7 more lightpaths (on average .5% improvement), in about one tenth of the computational times, except for instances NSF_1 and NSF_3. Indeed, while CG+ is much faster than CG, it requires more iterations (see the sum of the columns entitled #PP_iter), but as the pricing problem restricts the search to the shortest paths most of the time (indeed, 97% of the time on average), then computational times can be significantly reduced. Algorithm CG++ improves further the results of CG+ for all data instances from NSF_9 and on, increasing the number of granted lightpaths from 1 (NSF_9) to 14 (NSF_15) lightpaths. In addition, CG++ requires slightly more computational times than CG+, even if there are less wavelength configurations generated. This is due to the fact that PP$_{\text{PATH}}$ contains more paths and therefore requires longer computational times for its solution. However, the number of calls to PP$_{\text{LINK}}$ is less for CG++ than for CG+, this is as expected since PP$_{\text{LINK}}$ is richer in terms of the number of paths that are considered. Accuracy (ε) of the solution output by CG++ for the largest instance is 0.89%.

TABLE II
ALGORITHM COMPARATIVE PERFORMANCE ON THE NSFNET NETWORK

Data instance	$z^{\star}_{\text{LP}}$	CG [8]		CG+			CG++		
		$\tilde{z}_{\text{ILP}}$	#PP$_{\text{LINK}}$	$\tilde{z}_{\text{ILP}}$	#PP_iter		$\tilde{z}_{\text{ILP}}$	#PP_iter	
					PATH	LINK		PATH	LINK
NSF_0	430.0	412	85	419	158	6	415	127	2
NSF_1	519.0	499	116	503	223	8	501	119	1
NSF_2	575.0	553	106	556	161	6	558	129	2
NSF_3	706.0	681	128	688	288	9	686	141	1
NSF_4	751.0	729	132	734	225	7	730	132	2
NSF_5	838.0	818	141	821	272	9	818	151	1
NSF_6	919.0	898	108	902	191	6	899	130	1
NSF_7	995.0	974	126	977	180	6	976	148	2
NSF_8	1,078.0	1,055	105	1,059	165	6	1,059	125	1
NSF_9	1,159.0	1,137	106	1,138	136	2	1,143	137	2
NSF_10	1,242.0	1,218	109	1,221	189	7	1,222	108	1
NSF_11	1,338.0	1,314	112	1,316	183	6	1,321	155	2
NSF_12	1,422.0	1,398	122	1,403	170	7	1,405	163	3
NSF_13	1,501.0	1,474	106	1,479	179	6	1,485	162	2
NSF_14	1,578.0	1,548	126	1,550	207	5	1,563	172	2
NSF_15	1,656.0	1,638	88	1,627	152	6	1,641	189	4
NSF_16	1,694.0	1,667	97	1,667	159	7	1,676	119	2
NSF_17	1,775.0	1,748	104	1,748	182	7	1,758	109	1
NSF_18	1,849.0	1,825	109	1,827	171	7	1,835	127	2
NSF_19	1,924.0	1,898	92	1,900	167	6	1,907	145	2

TABLE III
COMPARISON OF COMPUTING TIMES - NSFNET NETWORK

Data instances	W	$z^{\star}_{\text{LP}}$	ε-optimal solutions					
			CG [8]		CG+		CG++	
			GoS	CPU (seconds)	GoS	CPU (seconds)	GoS	CPU (seconds)
NSF_0	30	430.0	94.5	20	96.1	3	95.2	25
NSF_1	35	519.0	93.1	32	93.8	33	93.5	57
NSF_2	35	575.0	87.5	30	88.0	7	88.3	9
NSF_3	45	706.0	92.4	43	93.4	24	93.1	76
NSF_4	45	751.0	88.1	39	88.8	7	88.3	27
NSF_5	50	838.0	88.8	41	89.1	8	88.8	48
NSF_6	55	919.0	88.5	31	88.9	3	88.6	55
NSF_7	60	995.0	88.1	33	88.4	3	88.3	6
NSF_8	65	1,078.0	88.1	26	88.4	3	88.4	7
NSF_9	70	1,159.0	88.8	29	88.9	2	89.3	5
NSF_10	75	1,242.0	88.7	26	88.9	3	89.0	4
NSF_11	80	1,338.0	89.1	31	89.2	3	89.6	4
NSF_12	85	1,422.0	89.0	35	89.3	3	89.4	4
NSF_13	90	1,501.0	88.7	27	89.0	3	89.4	5
NSF_14	95	1,578.0	88.4	35	88.5	3	89.2	3
NSF_15	100	1,656.0	88.8	22	88.2	3	88.9	3
NSF_16	100	1,694.0	86.2	24	86.2	3	86.7	3
NSF_17	105	1,775.0	86.2	27	86.2	3	86.7	3
NSF_18	110	1,849.0	86.9	28	87.0	3	87.4	3
NSF_19	115	1,924.0	86.5	24	86.6	3	86.9	2

TABLE IV
COMPARATIVE PERFORMANCE OF THE DIFFERENT ALGORITHMS ON THE USANET NETWORK

Data instance	$z^{\star}_{\text{LP}}$	CG [8]		CG+			CG++		
		$\tilde{z}_{\text{ILP}}$	#PP$_{\text{LINK}}$	$\tilde{z}_{\text{ILP}}$	#PP_iter		$\tilde{z}_{\text{ILP}}$	#PP_iter	
					PATH	LINK		PATH	LINK
USA_0	1,281.0	1,229	249	1,227	378	2	1,234	253	1
USA_1	1,507.0	1,453	292	1,437	329	1	1,448	295	1
USA_2	1,770.0	1,714	307	1,696	360	1	1,710	307	1
USA_3	2,035.0	1,943	326	1,951	334	1	1,964	384	2
USA_4	2,255.0	2,190	324	2,160	315	1	2,191	361	3
USA_5	2,516.0	2,418	292	2,435	307	1	2,461	337	2
USA_6	2,738.0	2,637	315	2,620	300	1	2,621	322	2
USA_7	2,886.0	2,722	271	2,735	288	1	2,765	357	3
USA_8	3,029.0	2,914	252	2,885	255	1	2,948	406	3
USA_9	3,188.0	3,027	286	3,012	256	1	3,063	354	2
USA_10	3,338.0	3,171	243	3,227	248	1	3,227	244	1

TABLE V
COMPARISON OF COMPUTING TIMES - USANET NETWORK

Data instances	W	$z^{\star}_{\text{LP}}$	ε-optimal solutions					
			CG [8]		CG+		CG++	
			GoS	CPU (seconds)	GoS	CPU (seconds)	GoS	CPU (seconds)
USA_0	75	1,281.0	92.0	663	91.8	35	92.4	70
USA_1	85	1,507.0	90.2	1,032	89.2	23	89.9	89
USA_2	100	1,770.0	90.8	1,225	89.9	34	90.6	83
USA_3	115	2,035.0	89.9	1,216	90.3	23	90.9	93
USA_4	125	2,255.0	90.4	1,211	89.2	21	90.5	91
USA_5	140	2,516.0	89.4	1,031	90.0	20	90.9	95
USA_6	150	2,738.0	88.7	1,191	88.2	19	88.2	92
USA_7		2,886.0	84.0	1,008	84.4	18	85.3	88
USA_8		3,029.0	83.0	890	82.2	15	84.0	80
USA_9		3,188.0	79.6	991	79.2	18	80.5	78
USA_10		3,338.0	77.6	878	79.0	13	79.0	72

TABLE VI
COMPARATIVE PERFORMANCE OF THE DIFFERENT ALGORITHMS ON THE GERMANY NETWORK

Data instances	$z^{\star}_{\text{LP}}$	CG [8]		CG+			CG++		
		$\tilde{z}_{\text{ILP}}$	#PP$_{\text{LINK}}$	$\tilde{z}_{\text{ILP}}$	#PP_iter		$\tilde{z}_{\text{ILP}}$	#PP_iter	
					PATH	LINK		PATH	LINK
GER_0	2,306.0	2,185	349	2,197	2,850	42	2,206	336	1
GER_1	2,960.0	2,791	437	2,840	3,543	104	2,861	432	2
GER_2	3,491.0	3,299	406	3,375	3,145	72	3,338	486	2
GER_3	3,911.0	3,700	501	3,771	3,362	61	3,782	527	1
GER_4	4,295.0	4,090	543	4,059	3,699	65	4,157	597	1
GER_5	4,663.0	4,472	731	4,502	6,253	119	4,507	1,443	11

2) NSFNET Results: We can make the following observations for the NSF network. Results of Tables II and III show that, on average, CG+ obtains better ε-values than

3) USANET Results: If we now turn to results obtained for USANET, conclusions differ a bit, most likely due to the

Data instance	$z_{\mathrm{LP}}^{\star}$	ε-optimal solutions					
		CG [8]		CG+		CG++	
		GoS	CPU	GoS	CPU	GoS	CPU
GER_0	2,306.0	92.4	6,045	92.9	887	93.3	90
GER_1	2,960.0	91.8	7,050	93.4	2,113	94.1	115
GER_2	3,491.0	89.9	8,589	92.0	1,522	91.0	101
GER_3	3,911.0	85.6	9,363	87.3	1,601	87.5	167
GER_4	4,295.0	82.0	9,400	81.4	1,379	83.3	252
GER_5	4,663.0	78.9	16,863	79.5	3,450	79.5	908

larger number of shortest paths per node pairs. In particular, CG+ improves on the values output by algorithm CG only for 4 over 11 instances, for an increased number of lightpaths ranging from 8 (in addition to 1,943 in USA_3) to 56 (in addition to 3,171 in USA_10). However, the differences between the computational times have widened, on average computational times for CG are 51 times longer than for CG+. It is interesting to observe that except for USA_0, we recourse to PP$_{\mathrm{LINK}}$ only one time, meaning that the solution provided by the wavelength configurations of PP$_{\mathrm{PATH}}$ account for most of the computational times, and explain the reduction of it in comparison with algorithm CG. As for NSFNET, CG++ improves on CG+ for most of the instances, while requiring to generate less wavelength configurations for the smaller USANET data instances, but more wavelength configurations for the larger USANET data instances. This can be explained by the larger number of shortest paths, and of second shortest paths, especially for some node pairs that have up to 45 different second shortest paths.

4) GERMANY Results: We next examine the results obtained for GERMANY. Except for one instance, GER_4, CG+ gets better solutions than CG, with an increased number of lightpaths ranging from 12 (in addition to 2,185 in GER_0) to 76 (in addition to 3,299 in GER_2). Of course, computational times of CG are higher than those of CG+, on average they are 5.5 times larger. As for previous networks, CG++ improves further the results of CG+, in all data instances except for GER_2, and always returns better results compare to CG. It is interesting to see that CG++ succeeds in obtaining better results in considerably shorter time than CG+, which itself performs dramatically faster than CG. In fact, the computational time of CG is on average 54.2 times longer than that of CG++, which is a significant improvement. This can be explained by the low number of recourse to PP$_{\mathrm{LINK}}$ in CG++ (from 1 to at most 11), that can make up for the larger solution time of its PP$_{\mathrm{PATH}}$ in comparison to that of CG+. Note that again, even the number of paths in P_{sd}^2 could be quite large, indeed, up to 70 for some node pairs in GERMANY.

VI. CONCLUSION

We obtained highly efficient algorithms for solving large scale RWA problems, which allow the exact solutions of data instances with up to 50 nodes and 150 wavelengths. It appears that the concept of limiting the search of wavelength configurations using shortest paths, or a limited number of

k-shortest paths is very promising in terms of enhancing the performance of the ε-optimal algorithms.

REFERENCES

[1] *Cisco Visual Networking Index: Forecast and Methodology, 20142019*, CISCO, May 2015.
[2] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, "Intertubes: A study of the us long-haul fiber-optic infrastructure," in *Conference of the ACM Special Interest Group on Data Communication -SIGCOMM*, London, United Kingdom, August 1721 2015.
[3] N. Nagatsu, S. Okamoto, and K. Sato, "Large scale photonic transport network design based on optical paths," in *IEEE Global Telecommunications Conference - GLOBECOM*, 1996, pp. 321–327.
[4] T. Noronha, M. Resende, and C. Ribeiro, "A biased random-key genetic algorithm for routing and wavelength assignment," *Journal of Global Optimization*, vol. 50, no. 3, pp. 503–518, 2006.
[5] A. Martins, C. Duhamel, P. Mahey, R. Saldanh, and M. C. de Souza, "Variable neighborhood descent with iterated local search for routing and wavelength assignment," *Computers & Operations Research*, vol. 39, pp. 2133–2141, 2012.
[6] R. Kogantia and D. Sidhu, "Analysis of routing and wavelength assignment in large WDM networks," *Procedia Computer Science*, vol. 34, pp. 71–78, 2014.
[7] T. Lee, K. Lee, and S. Park, "Optimal routing and wavelength assignment in WDM ring networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 2146–2154, October 2000.
[8] B. Jaumard, C. Meyer, and B. Thiongane, "On column generation formulations for the RWA problem," *Discrete Applied Mathematics*, vol. 157, pp. 1291–1308, March 2009.
[9] ——, "Comparison of ILP formulations for the RWA problem," *Optical Switching and Networking*, vol. 4, no. 3-4, pp. 157–172, November 2007.
[10] ——, "ILP formulations for the RWA problem for symmetrical systems," in *Handbook for Optimization in Telecommunications*, P. Pardalos and M. Resende, Eds. Kluwer, 2006, ch. 23, pp. 637–678.
[11] B. Jaumard, C. Meyer, and X. Yu, "When is wavelength conversion contributing to reducing the blocking rate ?" in *IEEE Global Telecommunications Conference - GLOBECOM*, vol. 4, 2005, pp. 2078 – 2083.
[12] ——, "How much wavelength conversion allows a reduction in the blocking rate ?" *Journal of Optical Networking*, vol. 5, no. 12, pp. 881–900, 2006.
[13] V. Chvatal, *Linear Programming*. Freeman, 1983.
[14] C. Barnhart, E. Johnson, G. Nemhauser, M. Savelsbergh, and P. Vance, "Branch-and-price: Column generation for solving huge integer programs," *Operations Research*, vol. 46, no. 3, pp. 316–329, 1998.
[15] J. Yen, "Finding the k shortest loopless paths in a network," *Management Science*, vol. 17, no. 11, pp. 712–716, 1971.
[16] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly, "SNDlib 1.0–Survivable Network Design Library," in *Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium*, April 2007, pp. 276–286.
[17] M. Batayneh, D. 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*, vol. 19, pp. 1304–1316, October 2011.