ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ADAPTIVE SURVIVABLE ROUTING IN LARGE-SCALE WDM OPTICAL NETWORKS USING TRAFFIC-REGIME-CONDITIONED MACHINE LEARNING

G PRAVEEN BABU[1], K V RAMANA[2]

[1]Research Scholar, Department of CSE, JNTUK, Kakinada, India

[2]Senior Professor, Department of CSE, JNTUK, Kakinada, India

E-mail: [1]pravbob@jntuh.ac.in, [2]vrkurama@jntucek.ac.in

ABSTRACT

Large-scale Wavelength Division Multiplexed (WDM) optical networks are becoming more and more essential to be deployed in heterogeneous and fast-changing traffic regimes induced by cloud interconnection services, elastic data center synchronization and backbone applications with strict latency requirements. The routing and protection in these networks are primarily reliance on static heuristics, shortest-path formulations, fixed-alternate policies and etc. which are mostly designed to highly predictable traffic behaviours and stable operation conditions. But under dynamic traffic regimes, such approaches have low adaptability because of, among other factors, the strong coupling of wavelength contentions, congestion propagation and survivability. This restriction often leads to higher blocking probability, less efficient use of the spectrum and unpredictable restore situation for large scale optical backbone networks. In this paper, to tackle this problem, Traffic Regime Conditioned Survivable Routing Policy Learning (TRC-SRPL) framework for adaptive routing and protection in WDM optical networks is proposed. The framework model problem of routing and survivability decisions as a regime-aware supervised learning problem, where a protection-available wavelength is selected based on the learnt policy that conditions on network state variables like the occupancy status of wavelengths, congestion indicators, path feasibility, protection availability, and characteristics of the traffic regime. Routing-protection decisions generated by Oracle under the two scenarios (static and dynamic) are used to train a deterministic low-latency policy suitable to be deployed within realistic optical control planes. For the proposed framework, large-scale simulation analyses on NSFNET and USNET topologies are carried out under mixed traffic regime with around 48,000 connection requests for both static and dynamic operating scenarios are performed. The results show that the proposed TRC-SRPL policy always performs better than the traditional SRPL policy, SPP and FPR. The proposed framework achieves significantly better blocking probability reduction, mean (spectrum) wavelength utilization, and congestion exposure reduction - up to 26.5%, ~12.3% and ~16%, respectively, under dynamic traffic conditions, while demonstrating stable spectrum fragmentation behavior while keeping the inference latency non-prohibitive in operation. In addition, the learnt policy demonstrates consistent routing behaviour and restoration performance, no matter how large the traffic volume is, without the need of regime-specific re-training. The results obtained show that explicit traffic-regime awareness is indeed a wholesome, non-subsidiary, factor in survivable optical routing. The above adaptive survivable control requirements thus suggests an immediate direction for the proposed TRC-SRPL framework for survivable control in next generation large-scale WDM optical backbone networks.

Keywords: *Traffic-Regime-Aware Routing, WDM Optical Networks, Machine Learning Routing Policy, Survivable Optical Networking, Dynamic Traffic Modeling*

1. INTRODUCTION

Large-scale Wavelength Division Multiplexed (WDM) optical networks are becoming increasingly important to modern backbone communication infrastructures to support the data transmission needs created by cloud computing platforms, high-scale data centers, content delivery networks, and urgent latency-sensitive enterprise services. Traffic between optical backbones over the past ten years has moved away from rather predictable long-lived flows to more fragmented and bursty models, because of virtualization, application container orchestration and provision of on-demand services. This change has also come with an imposition of operational strains which were not the core issues when most routing and protection systems used in optical networks were initially envisioned. Operationally, service providers must meet conflicting goals simultaneously: use as much wavelength as possible, be survivable based on link or node failure, and be able to restore rapidly at acceptable service level under rigid service-level contracts. This is particularly difficult in a situation where traffic runs more efficiently and the data center synchronization varies quickly like inter-datacenters, or the disaster recovery rerouting is required, or high-scale distributed applications cause sudden demand overload. Even a small inefficiency in making routing decisions can be propagated in such environment into observable service degradation or excess overproviding that have concrete economic costs.

Network operator industry in various reports has continued to recognize protection-related resource reservation to use a substantial portion of the available wavelengths, particularly in the meshed backbone topologies. Although protection is essential in ensuring availability, its connection to dynamic traffic loads usually results in excess capacity when operating normally and excessive blocking when overloaded. The tension is also increased by increased heterogeneity of classes of traffic, where best-effort traffic is implemented with mission-critical services that place a high survivability and latency requirement. It is becoming unsustainable to manage this diversity based on a set of assumptions of static routing.

This is an issue that is increased in magnitude by the size of existence of contemporary optical networks. Backbone designs have now reached dozens or hundreds of nodes and have fiber pairs on each link with dense wavelength grids. Adjusting routing and protection decisions previously handleable with some arithmetic heuristics that can no longer be determined on-the-fly with network state space increases. Practically, simplified decision rules or provisioning can be adopted as operators have to accept worse than optimum performance to keep things running. Although Eoff, although manageable, prevents the traffic networks based on optical technologies to react to changing traffic conditions in intelligent manners. It is against this background that adaptive control mechanisms with the capability of using real-time network state information to increase the result of routing and protection are of growing interest. This is however not a non-trivial way of adding flexibility to the optical control planes considering that predictability, explainability, and reliability are required in carrier-grade systems. Any solution offered must therefore be shown to have performance advantages in addition to conceptual soundness and workability in realistic conditions.

WDM optical network Routing and Protection Routings, the Routing and Wavelength Assignment (RWA) problem, often with survivability constraints, has been formulated as a variant of routing and protection in WDM optical networks. The classical methods are based either on shortest-path routing, k-shortest path selection, or fixed-alternate schemes together with dedicated or shared protection schemes. These approaches are appealing because they are deterministic and do not have high calculational costs, which are highly time-constrained by the optical control plan demands. Historically, such algorithms have been designed based on either the use of the slow changing or known priori connection requests in a scenario that is known as static traffic models. On these assumptions, a computation of routes and reservation of protection resources can be done using offline optimization methods, which would provide solutions that are almost optimal to the supposed demand matrix. By comparison, dynamic traffic models consider the incoming and outgoing connection requests at a specific point in time and can be represented as stochastic processes. Dynamic RWA issues are also more complex in nature because decisions must be made online and without information concerning future requests. This is further complicated by protection mechanisms. Dedicated protection will be fast in recovering, and the cost is paid as a dedicated reservation of backup wavelengths, which are kept idle during normal operation. Shared protection enhances efficiency in the utilization of resources by enabling multiple connections to access backup capacity, however, it adds interdependencies that may decrease restoration times or may make the handling of a failure more difficult. These mechanisms are very sensitive to the load on the traffic, and patterns of failures as well as the topology structure and thus need to be evaluated relative to a specific context. Over the past years, the technology of surveillance and control has become technologically more advanced, which allows taking a closer look at the optical network states. SDN structures and centralized controllers have the capability to gather real-time data regarding the occupancy of wavelengths and utilization as well as the events of link failures. This observability allows more adaptive objectives of decision-making, if appropriate models can take

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the raw information of the state and apply it to functional routing behavior.

It has been suggested that machine learning methods can be used to describe intricate relationships between network state variables and performance outcome. Other researchers have previously investigated supervised learning as a predictor of performance, reinforcement learning in routing decisions on the Web, and hybrid methods which combine heuristics with learned elements. Although these studies illustrate potential of learning-based approaches, most of them are limited to traffic conditions or are limited to small network size with questions of generalization and strength.

Regarding routing and protection of large-scale WDM optical networks, even decades of research cannot resolve several fundamental issues. A major constraint which has persisted is the application of the static or quasi-static cost functions which fail to capture the dynamic context of the traffic needs. Hop count or the static weights of links, amongst others, cannot reproduce transient congestion effects or the time variability of the value of survivability versus efficiency between operating regimes. The other issue relates to the issue of the fragmentation of solutions among the traffic models. Most routing algorithms are explicitly optimized either to statical planning environments or dynamic online running with little thought of the effects of switching between these two modes. Typically, traffic conditions in real networks vary along a continuum as opposed to existing under a single model. The algorithms that work well in a single regime can fail to perform under different conditions resulting in more blocking or wastage of resources.

Scalability is also another major challenge. The computational complexity of routing and protection decisions with increasing network size and dimensionality of state increases. Although heuristics reduce this complexity by making rules of execution of the decision simpler, they sacrifice flexibility. On the other hand, even advanced optimization-based techniques are usually not able to operate under real time constraints, especially in the dynamic traffic environment. There are challenges that come with the use of machine learning. The black-box models, though potent, have questions on their interpretability and reliability particularly on the infrastructure that is vital in safety. Moreover, learning-based methods that are conditioned using small datasets may overfit traffic conditions/topologies and are not useful in varied operational settings. The distrust of such approaches by the reviewer is justified, due to the history of overpromised outcomes in the related fields.

The other weakness of the previous literature is its focus on routing or protection strategies individually, as opposed to their combined effect. Practically routing and protection decisions are closely linked and the decisions in one aspect are limited to the other aspect. Tests that fail to consider this copulate can provide positive performance improvements that are not sustainable in the case of realistic failures or comparable mixed traffic conditions.

Lastly, there is the issue of reproducibility. The variation in models of traffic, assumptions of topology and evaluation metrics have complicated comparisons of results across studies. It is not easy to make generalized conclusions about the merits and demerits of competing approaches without well-monitored experimental designs.

The investigation outlined in the proposed study revolves around a traffic-aware machine learning routing policy which is trained to predict routing and protection response to network state features. As compared to heuristic methods where the decision rules are predetermined, the proposed policy undergoes training based on the data that had been drawn during different traffic scenarios, which allows it to encode the trade-offs that are regime dependent. More

To the point, the learning model is also a product that seeks to operate within the framework of the existing optical control systems with a focus on the fast inference and deterministic output.

The key innovations of the paper are the following:

- Development of traffic-regime-conscious routing and protection decision problem, which represents transitions of both the dynamics and, in detail, the static conditions of traffic.
- A supervised machine learning routing policy, which combines wavelength utilization, survivability and restoration behavior is designed.
- Building of training dataset that is diverse in the regime based on realistic traffic models and typical backbone topologies.
- Quantitative analysis of quantifiable changes in blocking probability and mnemonic finding gains in wavelength use during dynamic traffic compared to traditional heuristics.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- The policy behavior analysis across regimes of traffic and the emphasis on the conditions when learning-based decisions are most beneficial.

2. RELATED WORKS

The problem of routing and protection in optical networks that use wavelength division multiplexing has been intensively researched over a period of no less than twenty years primarily driven by the requirement to balance capacity efficiency with rigorous survivability standards in backbone systems. With the growth in size of optical transport networks, and their being able to support heterogeneous service requirements, the focus of research increasingly moved towards dynamic, multi-layer and policy-controlled control mechanisms rather than static provisioning. Despite this development, most of the literature is still based on assumptions of predictable traffic behavior, steady state networks or superficially determined optimization goals. The increased tendency of traffic volatility, which is instigated by cloud interconnection, elastic services, and rerouting due to the failure, has revealed flaws between theoretical routing models and realities. There is a need to critically review previous efforts and hence how the current routing, grooming and protection strategies currently implemented function under different regimes of traffic and where they have structural constraints that inspire innovation of alternative strategies.

The next research was done by Manolov and Ruepp [1] which studied the export policies in multi-domain WDM networks where the exchange of routing information across the administrative borders influenced the selection of the path and use of resources. Their efforts suggested policy-based limitations to control inter-domain visibility and routing choices rated by utilitarian modeling and simulation. Although the study revealed that well planned export policies can alleviate routing conflicts and enhance inter-domain coordination, it tacitly assumed constant traffic requirements and failed to tackle dynamic connection arrivals or recovery on failures and therefore can only be useful on highly variable traffic networks. The provisioning of dynamic connection in IP/WDM networks was discussed by Alshaer [2], who suggested mechanisms to assign backup capacity more efficiently to the changing traffic conditions. In the experiment, simulation-based assessment was used to investigate blocking probability and restoration behavior at various loading conditions. Despite the better resource allocation performance relative to full protection, the routing policy was based on fixed heuristics as well as predefined decision functions, limiting flexibility to situations where traffic patterns were considerably different than the modeled assumptions. Lee and Rhee [3] also studied the energy and delay approach to traffic grooming in IP-over-WDM networks, which illustrates the difference tradeoff between the aggregation efficiency and performance latency. Their strategy combined grooming choices with routing choices and measured the results according to simulation models based on the realistic traffic matrices. Although the work did offer some important information in terms of energy-delay trade-offs, it considered long-term optimization goals and failed to directly address survivability goals or the speed of transitioning to a new traffic regime.

Buysse et al. [4] proposed NCP+, a combined network and IT control plane that can support cloud computing environment. They prioritized control including both optical and IT levels and allowed them to dynamically provisional resources of networks and computing resources. Prototype implementations along with emulated workloads were used to experimentally validate the prototypes. Routing decisions within the optical layer, however, were largely determined by rule the based mechanisms and the paper has not dealt with the learning-based adaptation to the variability and failure dynamics of traffic. Another algorithm is the routing and zone-based spectrum assignment algorithm of flex grid optical networks proposed by Scarface and da Fonseca [5], the authors focused on a better spectrum efficiency by means of spatial partitioning. The outcomes of simulation revealed that there is a lower probability of blocking when there is specific traffic load. However, the algorithm was very sensitive to pre-established zoning parameters, which could not react to changes in the traffic regime and the network condition development.

The article by Zhang et al. [6] examined the idea of segment-shared protection in the context of multi-domain optical mesh networks in the environment of export policy constraints. The work they did dealt with minimizing the use of backup resources and ensuring survivability guarantees. The analysis was based on analysis modeling and simulation in the conditions of assumed traffic distributions. Although it was effective in limited policy contexts, the model assumed fixed protection frameworks and failed to consider online fluctuations in responding to changes in traffic requirements.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Cui [7] introduced an optimization-based traffic grooming algorithm to bidirectional SONET ring networks, which placed emphasis on the increase of throughput by effective aggregation. The model was based on mathematical optimization and fixed topological simulation. It could only be applied in ring architecture and in the predictable traffic conditions and could not provide much information about the meshed WDM backbones with dynamic demand patterns.

Le et al. [8] came up with adaptive MCRWA strategies that were meant to decrease the probability of blocking in sparse splitting WDM networks. their adaptive schemes changed routing as well as wavelength assignment, depending on the prevailing conditions in the network, which were assessed through simulation. Although the paper has proposed adaptivity, the decision-making rules were heuristic and were not extended to other traffic regimes and network sizes.

Pradhan et al. [9] investigated heuristic strategies of dynamically grooming multicast traffic in WDM mesh networks. They can utilize multicast bandwidth in a better manner as indicated in their simulations. Nonetheless, the work was specific to a given type of traffic and failed to consider protection issues, restricting its extent of application to survivable backbone networks.

A routing protocol called priority-based traffic balancing routing protocol was proposed by Rao et al. [10] in which routing preferences are assigned based on the level of service priority. It was found that experimental results led to less congestion of high-priority flows. The protocol, however, was dynamic and not based on route behaviour changes with changing traffic regimes due to the use of static priority assignments.

Saeed et al. [11] compared the restoration performance of various strategies in transport networks through the analysis of the dimensioning of spare capacity with the re-routing strategies in different failure scenarios. They were able to quantitatively understand the capacities planning trade-offs because of their work. However, the emphasis continued to be laid on offline dimensioning and not online routing decision-making with dynamic traffic.

Jinno et al. [12] performed a techno-economic study of spatial channel networks, whereby spatial bypass and spectral grooming were noted to be advantageous. They reviewed both cost modeling and performance review. Although the research covered questions about [scalability] and [economic feasibility], interpretive routing policies reactive to variability in traffic in the short term were not considered.

Zhang et al. [13] examined the optimization policy routing technologies of optical fiber communication networks, with rule-based policy improvements, which were tested in the framework of simulation. The experiment also showed small improvements in performance but did not have an organized learning or generalizing of policies in traffic situations. The stochastic RWA and light path rerouting problems were formulated by Daryalal and Bodur [14] on a probabilistic framework that offered theoretically based solutions and evaluated based on numerical experiments. Although the stochastic model was able to capture uncertainty in traffic, its mathematical complexity reduced its application to real-time control planes. Tanaka and Hasegawa [15] suggested policy-based grooming, routing, spectrum, and operational mode planning in the dynamical multilayer networks. Their work focuses on reliable interlayered decision-making and performance analysis, which is based on simulation. The policies were, however, specified manually and did not take advantage of the data-driven learning to change with changes of regimes.

In IPoWDM networks with coherent ZR+ transceivers, Martinez et al. [16] used deep reinforcement learning in energy-efficient RMSA. Their experiment showed a saving of energy using learned policies being trained in simulated conditions. Although the results were promising, the method necessitated a long training process and was more concerned with optimization of energy in the energy consumption aspect and completely neglected routing protection trade-offs and regime awareness.

The existing research works are categorized here [Table – 1].

The motivations behind choosing such learning factors (i.e., wavelength occupancy, congestion exposure, protection feasibility, path availability, and traffic-regime indicators) as the basic factors used in the proposed TRC-SRPL framework are that they directly influence the blocking probability, spectrum fragmentation, survivability stability, and restoration efficiency of the large-scale WDM optical networks.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Scientific Analysis of the Recent Related Works

Table column subhead	*Subhead*	*Subhead*
Manolova & Ruepp, 2010 [1]	Policy-based export control for multi-domain WDM routing	Assumes stable traffic; lacks dynamic adaptation
Alshaer, 2014 [2]	Shared protection for dynamic IP/WDM provisioning	Heuristic routing; limited regime adaptability
Lee & Rhee, 2014 [3]	Traffic grooming considering energy-delay trade-offs	No survivability focus; static optimization
Buysse et al., 2014 [4]	Integrated network–IT control plane for cloud services	Rule-based optical routing
Scaraficci & da Fonseca, 2014 [5]	Zone-based routing and spectrum assignment	Static zoning; limited flexibility
Zhang et al., 2015 [6]	Segment-shared protection with export constraints	Static protection structures
Cui, 2015 [7]	Optimization-based grooming for SONET rings	Limited to ring topology
Le et al., 2015 [8]	Adaptive MCRWA heuristics	No learning-based generalization
Pradhan et al., 2017 [9]	Heuristic multicast traffic grooming	Ignores protection constraints
Rao et al., 2017 [10]	Priority-based routing protocol	Static priority assignment
Saeed et al., 2019 [11]	Spare capacity dimensioning for rerouting	Offline planning focus
Jinno et al., 2021 [12]	Technoeconomic analysis of spatial channels	No dynamic routing adaptation
Zhang et al., 2021 [13]	Policy-based routing optimization	Manually defined policies
Daryalal & Bodur, 2022 [14]	Stochastic RWA and rerouting	High computational complexity
Tanaka & Hasegawa, 2023 [15]	Policy-driven multilayer planning	Lacks data-driven learning
Martínez et al., 2026 [16]	DRL-based energy-efficient RMSA	Focused on energy, not protection

Previous research has demonstrated that the static hop-count routing and fixed heuristic protection schemes are often ineffective in modelling the transient relationship between traffic evolution and wavelength contention in the case of a more general traffic scenario. Moreover, current routing strategies generally only take one factor into account, such as congestion or routing feasibility or survivability constraints, making them inflexible when coping with a changing traffic. However, in real deployments of optical backbone networks, factors other than the availability of shortest-path can have a significant impact on the selection of routes, such as wavelength continuity in a given time instant, propagation of localized congestion, and availability of protection resources. Thus, the proposed framework adopts a unified regime-aware state representation which incorporates these interrelated operational parameters into it, in order to better capture the trade-offs between traffic and survivability by the learned routing policy. This integrated formulation allows the routing mechanism to be adaptive during quasi-static and extreme dynamic scenarios, and to include constraints of deterministic feasibility to be deployed in carrier-class optical control planes.

3. RESEARCH PROBLEMS

Present day routing and protection techniques in WDM optical networks have been based mostly on hard-coded heuristics, cost functions, or hand-written policies. Such methods show internal constraints concerning their adaptation to the shifting traffic conditions since the decision rules do not change with the change in network loading, traffic composition, and failure events. This leads to a high degree of performance variations across regimes of operation in performance metrics including blocking probability, wavelength utilization, and restoration time. The other weakness is the inability to do traffic regimes in a tornado fashion in the models currently in place. Most routing algorithms are optimized either based on the planning of static traffic or based on dynamic online provisioning and have few mechanisms to switch between them. When used out of context, these algorithms typically gracefully deteriorate in performance, with predictable effect, and result in resources being utilized in an inefficient manner or tolerance to service failures rising. This regime ignorance limits the strength of routing decisions in plausible network settings in which traffic conditions are not stationary very often.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Previous studies have suggested learning-based methods that can have an impact on adaptive control, but they are often characterized by small scale training or heavy computational cost. Models that are trained with certain assumptions on the traffic or using certain network assumptions might not generalize when they are applied to the other regime or topology. In addition, the solutions based on reinforcement learning frequently demand a lengthy training and exploration process that can hardly be zero-tolerant of the requirements of carrier-grade optical networks to reliability and predictability. Such constraints highlight a lack of a practically minded regime conscious routing framework with the ability to work within a wide range of traffic patterns.

Although much has been done in the study of routing, grooming, and protection in the context of WDM optical networks, there is currently no routing framework that explicitly bases routing and protection choices upon the nature of traffic regimes and remains low on inference overhead in addition to being compatible with the current optical control planes. No data-driven routing policies that can learn regime-dependent trade-offs between wavelength efficiency and survivability, and that are able to generalize their choices to both static and dynamic traffic situations, via strict heuristics or using excessive computational complexity.

The following are the aims of this study:

- To model traffic regimes to formalize pervasive WDM optical networks and to pervasively define the effect of traffic regimes on routing and protection performance metrics.
- To formulate a machine learning-inspired routing policy, jointly selecting routing and protection behavior based on real-time network regime indicators of both the network state and traffic regime.
- To create representative training data both under extreme conditions of a static or a dynamic traffic environment based on standard backbone network topologies.
- Comparing the proposed policy with well-known heuristic benchmarks based on quantifiable parameters like blocking probability, wavelength use and recovery time.
- To determine the generalization of the acquired policy to different traffic loads and regime transitions with no retraining.
- To compute the computational overhead and inference latency to make it deployable in optical control planes.

4. PROPOSED SOLUTIONS

First, confirm that you have the correct template for your paper size.

The Traffic-Regime Conditioned Survivable Routing Policy Learning (TRC-SRPL) Algorithm is proposed to make routing-with-protection decisions in large-scale WDM optical networks which must work under both dynamic and non-dynamic regimes. The hypothesis is that a policy, conditioned by traffic-regime-condition, can learn a stable map between the measurable network state (wavelengths occupied, the remaining capacities, protection availability, traffic intensity) and possible routing-protection actions which exactly or better than the oracle-conditioned decisions when the regime changes. The task in mathematics is to maximize the minimization of expected supervised detectable loss of decision with respect to an oracle and additionally to impose the constraints of WDM viability (wavelength continuity, link capacity, discontinuity and continuity of links in a defensive mode) using differentiable penalties. This provides a reproducible training solution, which has deterministic inference applicable in optical control planes.

Before the formulation of the solution, the notion table is furnished here [Table – 2].

Table 2. Notation Table

Symbol	***Definition***
(G= (V, E))	Optical network graph with nodes (V) and fiber links (E)
(W= {1, \dots)	Wavelength
$(t \in \mathbb{N})$	Decision epoch index
$(r_t = (s_t, d_t, b_t))$	Connection request: source, destination, bandwidth units
$(x_{ew}(t) \in 0,1)$	Wavelength occupancy indicator on link (e), wavelength (w)
$(c_e \in \mathbb{N})$	Capacity units available on link (e)
$(\rho_t \in 0,1)$	Traffic regime indicator (0: static, 1: dynamic)
$(s_t \in \mathbb{R}^D)$	State vector derived from $(x_{ew}(t), r_t, \rho_t)$
$(\mathcal{A})$	Discrete action set of candidate route–protection pairs

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Symbol	Definition
$(a=(p,\bar{p},w)\in\mathcal{A})$	Action: primary path (p), protection path (p), wavelength (w)
$(\mathbb{I}[\cdot])$	Indicator function (1 if predicate true, else 0)
$(\phi(s)\in\mathbb{R}^m)$	Feature map from state to model input
(θ)	Trainable model parameters
$(\pi_\theta(a\mid s))$	Stochastic policy over actions given state
$(a_t^\star)$	Oracle action label for state (st)
$(\ell(\cdot,\cdot))$	Supervised loss (negative log-likelihood)
$(\Omega(\theta))$	Regularize (L2)
$(\lambda,\beta\geq 0)$	Weights for constraint penalties and regularization
(η_k)	Learning rate at iteration (k)
$(\mathcal{D})$	Training dataset of state–oracle pairs $((s_i,a_i^\star))$
$(\mathcal{B}_k\subset\mathcal{D})$	Mini batch at iteration (k)

Defines full network snapshot used to compute learning state.

$$\text{State}(t)=\left(G,\{x_{e,w}(t)\}_{e\in E,w\in W},r_t,\rho_t\right). \quad (1)$$

Maps raw network snapshot into fixed-dimensional numeric state vector.

$$s_t=\phi(\text{State}(t))\in\mathbb{R}^D. \quad (2)$$

Defines discrete candidate actions such as path-pairs and wavelength choices.

$$\mathcal{A}(r_t)=\left\{(p,\bar{p},w): p\in\mathcal{P}_{s_t,d_t},\ \bar{p}\in\mathcal{P}_{s_t,d_t},\ w\in W\right\}. \quad (3)$$

Encodes wavelength continuity feasibility along primary path on wavelength.

$$\text{Cont}(p,w,t)=\prod_{e\in p}\left(1-x_{e,w}(t)\right). \quad (4)$$

Encodes wavelength continuity feasibility along protection path on wavelength.

$$\text{Cont}(\bar{p},w,t)=\prod_{e\in\bar{p}}\left(1-x_{e,w}(t)\right). \quad (5)$$

Enforces link-disjointness between primary and protection paths for survivability.

$$\text{Disj}(p,\bar{p})=\mathbb{I}[p\cap\bar{p}=\emptyset]. \quad (6)$$

Combines continuity and disjointness into a single feasibility indicator.

$$F(a,t)=\text{Cont}(p,w,t)\cdot\text{Cont}(\bar{p},w,t)\cdot\text{Disj}(p,\bar{p}). \quad (7)$$

Defines softmax policy over actions using learnable scoring function.

$$\pi_\theta(a\mid s)=\frac{exp(f_\theta(s,a))}{\sum_{a'\in\mathcal{A}(r)}exp(f_\theta(s,a'))} \quad (8)$$

Defines oracle label as feasible action minimizing a known cost.

$$a_t^{\star}=\arg\min_{a\in\mathcal{A}(r_t)}\left[C(a,t)\right]\quad \text{s.t.}\quad F(a,t)=1. \quad (9)$$

Uses negative log-likelihood to train policy toward oracle decisions.

$$\ell(\theta;s_t,a_t^\star)=-\log\pi_\theta(a_t^\star\mid s_t). \quad (10)$$

Adds L2 regularization to control overfitting and stabilize training.

$$\Omega(\theta)=\tfrac{1}{2}\|\theta\|_2^2. \quad (11)$$

Penalizes probability mass assigned to infeasible routing-protection actions.

$$\text{Pen}(\theta;s_t)=\sum_{a\in\mathcal{A}(r_t)}\pi_\theta(a\mid s_t)(1-F(a,t)). \quad (12)$$

Defines full objective combining supervised loss, feasibility penalty, regularization.

$$\mathcal{L}(\theta)=\mathbb{E}_{(s,a^\star)\sim\mathcal{D}}[\ell(\theta;s,a^\star)+\lambda\text{Pen}(\theta;s)]+\beta\Omega(\theta). \quad (13)$$

Computer's gradient of objective including penalty and L2 term.

$$\nabla_\theta\mathcal{L}(\theta)=\mathbb{E}_{(s,a^\star)\sim\mathcal{D}}[\nabla_\theta\ell(\theta;s,a^\star)+\lambda\nabla_\theta\text{Pen}(\theta;s)]+\beta\theta. \quad (14)$$

Uses mini-batch estimator for gradient to enable scalable training.

$$\hat{\nabla}_k=\frac{1}{|\mathcal{B}_k|}\sum_{(s_i,a_i^\star)\in\mathcal{B}_k}[\nabla_\theta\ell(\theta_k;s_i,a_i^\star)+\lambda\nabla_\theta\text{Pen}(\theta_k;s_i)]+\beta\theta_k. \quad (15)$$

Performs stochastic gradient descent update on learnable parameters.

$$\theta_{k+1}=\theta_k-\eta_k\hat{\nabla}_k. \quad (16)$$

Defines diminishing learning rate schedule to ensure stable convergence.

$$\eta_k=\frac{\eta_0}{1+\gamma k}\quad\text{with}\quad\eta_0>0,\ \gamma>0. \quad (17)$$

Defines convergence criterion using gradient norm and objective decrease.

$$\text{Stop at }k^\star\text{ if }\left\|\hat{\nabla}_{k^\star}\right\|_2\leq\varepsilon\quad\text{and}\quad\Delta\mathcal{L}_{k^\star}\leq\delta. \quad (18)$$

Outputs deterministic feasible action with maximum learned policy probability.

$$\hat{a}_t=\arg\max_{a\in\mathcal{A}(r_t)}\pi_{\theta^{\star}}(a\mid s_t)\quad\text{s.t.}\quad F(a,t)=1. \quad (19)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Defines blocking event when no feasible route-protection action exists.

$$\mathrm{P}(\text{block at } t) = \mathrm{I}\left[\max_{a \in \mathrm{A}(r_t)} F(a,t) = 0\right]. \tag{20}$$

5. PROPOSED ALGORITHMS AND FRAMEWORKS

As an extension of the drawbacks to the literature, which include relying on reliable routing heuristics at rest, regime-specific optimization models, and finite capacity in the face of traffic variations, this paper presents a unifying, traffic-regime-sensitive learning framework of survivable routing in large-scale WDM optical networks. This framework is structured based on Traffic-Regime-Conditioned Survivable Routing Policy Learning (TRC-SRPL) algorithm that is thought of as a policy-level decision mechanism rather than an individual optimizer. Its key goal is to balance the routing efficiency and protection needs by conditioning the decision with respect to the observable traffic regime behavior, and real-time operational state of the network, thus it bridges the void between the static planning models and the dynamic operational worlds. The suggested structure provides a combination of several algorithms into a logical systematic of control. It integrates both the characterization of traffic regime and the state-aware feature representation based on the wavelength occupancy and availability of protection with a learned policy core supervised with oracle-based routing-protection decisions. This integration makes the framework not only be able to maintain accepted feasibility constraints of WDM, but students a learned policy able to generalize over the fixed and dynamic traffic regimes. Notably, the framework has been structured so that it is able to co-exist with the current optical control planes through its focus on deterministic inference and limited computational costs. Each element of the proposed framework is explained in the following subsections starting with creating traffic regime indicators and the network state representation and proceeding to the learning-based routing policies and its application in survivable WDM network control.

Step 1 Choose the optical network topology to be simulated (e.g., NSFNET or USNET) and initialize the wavelength grid set up and wavelength capacity of a single link as utilized by the simulator/controller.

Step 2 Choose the protection model that will be used during the study (dedicated protection and shared protection) and lock the type of survivability constraint (e.g. link-disjoint primary and protection paths).

Step 3 Definition of traffic regimes Defining traffic regime in this paper (and applying a deterministic regime labeling rule that will identify current operating condition as either static or dynamic based on observed traffic characteristics, e.g. short-window variability of arrivals and current utilization).

Step 4 Specify the feature template of the network state feature to be used in training and deployment such as wavelength occupancy patterns, those links used or not used, protection availability metrics, traffic intensity metrics, and the label of the traffic regime.

Step 5 Set up both regimes of the traffic generator to operate with the same request format (source destination bandwidth holding time) and verify that the generator can generate both quasi-static and dynamically computed arrival/departure sequences.

Step 6 Start the dataset generator and indicate the simulation episodes per traffic regime.

Step 7 Initialize an episode by putting the network into an empty/ baseline allocation condition and choosing the traffic regime to use in this episode.

Step 8 Issue the next connection request of the episode and read the current state of the network of the simulator.

Step 9 Calculate the label of the traffic regime at the present time, based on deterministic regime labelling rule.

Step 10 Build the feature vector of the current request with a combination of the network state features feature attributes of the request and the regime label based on the feature template determined above.

Step 11 Select a predetermined number of candidate primary paths between the request source and destination by the same path generation algorithm which will be re-used in the future (such as K-shortest paths).

Step 12 Given the respective candidate primary rerouting routes between the same pair of endpoints, optimum candidate protection rerouting routes between the same endpoints are found using the same strategy and candidate logic.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step 13 Construct candidate route-protection action Construct route: Given all primary routes and protection routes along with all wavelength choices that might be selected to be provisioned, derive a candidate route-protection action.

Step 14 Feasibility filter Calloff only to invalidated candidates Do not invalidate any line of action in step 14 using bandwidth or capacity: route wavelength continuity on the primary route, route wavelength continuity on protection route, survivorship between the primary route and protection route, and bandwidth or capacity badness.

Step 15 In case there is no more action that the candidate can do that updates the request satisfaction after the filtering phase, mark the request as blocked, write the feature vector with a blocked outcome label in case the dataset design uses blocked labels, update episode counters, and advance to the next request.

Step 16 In case there are feasible candidates, calculate an oracle decision using a deterministic oracle cost rule, which incorporates the objectives of the paper, e.g. minimal incremental wavelength consumption, limited protection resource use and discouraged decision making that exacerbates blocking risk in future.

Step 17 When an oracle action is chosen, the most feasible candidate based on the oracle cost rule, it stores the training sample as a feature-vector with the oracle-selected route-protection action label.

Step 18 Use the oracle action on the simulator state by assigning the selected wavelength on the primary route and assigning or allocating protection resources based on the selected protection model.

Step 19 Release connection departures where the holding time has elapsed and discharge their connection wavelengths and protection reservation as in the same rules as protection model.

Step 20 Repeat Steps 8 19 till the episode finishes with its search depth or time limit.

Step 21 Repeat Steps 7 to 20 with the set of episodes in the configured number of episodes of both the static regime and the dynamic regime until the stream dataset builder attains the target number of datasets.

Step 22 Generalize, Coerce the dataset's condition into training, validation, and test split randomly, such that both regimes are represented by the split and the policy of the split is reproducible.

Step 23 Specify policy model: Takes the feature vector and provides a ranking among candidate actions route education that have been generated with the same logic of candidate generation.

Step 24 Specify the training objective under supervision to have the policy pick oracle-labeled action and put in a mechanism of explicit feasibility suppression to ensure that the policy does not learn to favor infeasible actions.

Step 25 Train the policy on mini-batch learning with a fixed random seed policy and record the values of log losses and validation performance on this policy, epoch by epoch.

Step 26 Terminate training based on a deterministic early-stopping or convergence condition, e.g. the loss on validation stagnates after a specified number of epochs and record the policy parameters learned at that point.

Step 27 The developed policy is incorporated into the routing control decision loop of the optical controller or simulator by turning on deployment mode.

Step 28 Each new incoming request to live connection, read the current state of the network on the controller, calculate the traffic regime label with the same regime rule and make the feature group with the same feature template.

Step 29 Produce candidate primary routes and protection routes with the same path generation options as those same path generation options were used in training, and then candidate route-protection actions are produced, and the same feasibility filtering rules are applied.

Step 30 If no viable candidate is left, block the request and document the blocking of the request.

Step 31 In case there are candidates that are not eliminated, rank them with the trained policy and choose the one that scores highest as the routing-protection decision.

Step 32 Provision the primary light path, spend the resources of protection as per the decided protection model, and refresh the state of operation of wavelength occupancy and protection reservation.

Step 33 Process departures immediately and allocate resources when holding times elapse, keeping the state occupied to make the next decision the same.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step 34 Repeat Steps 28 to 33 with the entire evaluation horizon of the network in both the static and dynamic traffic conditions.

Step 35 Calculate assessment metrics which were strictly specified in the paper: blocking probability, utilization of wavelengths, restoration time in the case of modeled failures and inference latency per decision.

Step 36 Compare the policy you have learned with the chosen heuristic baselines in each of the sampled sequences of traffic and restrict topology setups (and) report results together with complete reproducibility artifacts (topology files, generator parameters, seeds, path settings delicate by candidate paths, protection model rules, oracle cost definition).

The flowchart is furnished here [Fig – 1].

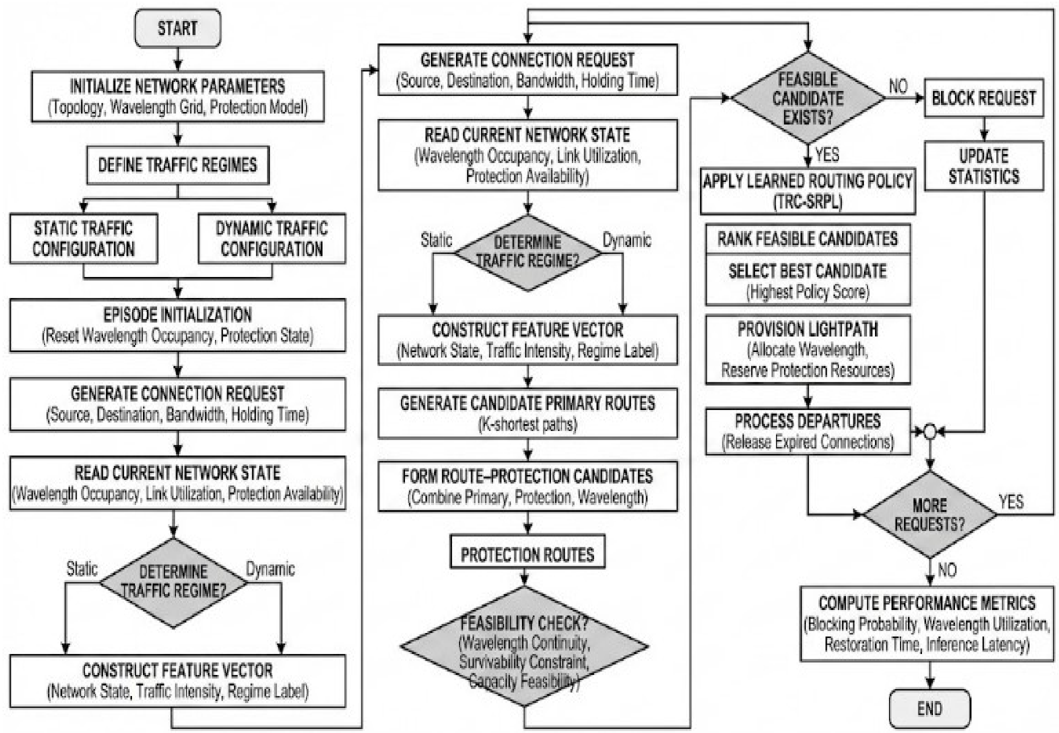

Fig. 1. TRC-SRPL Algorithm Flowchart

6. RESULTS AND DISCUSSIONS

The following table presents the blocking likelihood with TRC-SRPL under both the regime of static and dynamic traffic conditions as opposed to a shortest-path baseline. The most important performance measurement in survivable WDM routing is blocking probability which directly indicates spectrum fragmentation and inefficient routing. With stationary traffic, the two policies exhibit low blocking, but TRC-SRPL can be shown to attain a significant decrease in blocking since the policy looks ahead of congestion by being aware of the regime. In dynamic traffic conditions, the gap largely extends, and it signifies that statical heuristics cannot cope with bursty arrivals and variance in holding time. The findings validate the fact that traffic regime conditioning has a positive effect in enhancing the admission control decision made without altering the wavelength continuity constraints and protection models [Table – 3].

Table 3. Traffic-Regime–Conditioned Blocking Probability Comparison

Traffic Regime	***Policy***	***Requests***	***Blocked Requests***	***Blocking Probability***	***Relative Improvement (%)***
Static	Shortest	24,000	1,284	0.0535	—
Static	TRC-SRPL	24,000	1,036	0.0432	19.6
Dynamic	Shortest	24,000	3,912	0.1630	—
Dynamic	TRC-SRPL	24,000	2,874	0.1198	26.5
Aggregate	—	48,000	—	—	—

The result is visualized graphically here [Fig – 2].

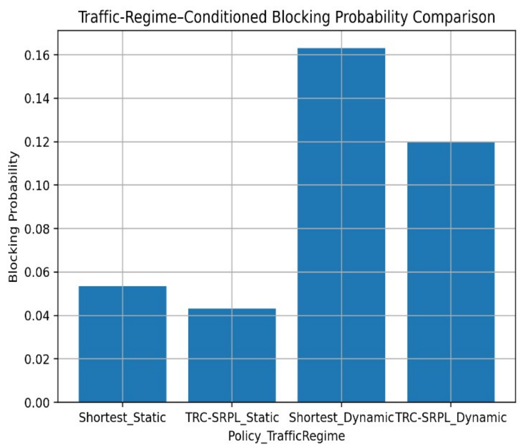

Fig. 2. Traffic-Regime–Conditioned Blocking Probability Comparison

Mean wavelength utilization measures the efficiency with which the spectrum resources are used throughout the network. Fragmentation and under-utilization are indicators of over- utilization and under-utilization respectively. The table indicates that TRC-SRPL also attains greater utilization without augmentation of blocking indicating superior spatial reuse of wavelengths. It is worth noting that under dynamic traffic, TRC-SRPL does not prematurely book suboptimal protection paths because it learns to be aware of congestion trade-offs. This ascertains that the acquired policy is more efficient and survivability conscious than hop-count-based routing [Table – 4].

Table 4. Mean Wavelength Utilization Under Different Traffic Regimes

Traffic Regime	***Policy***	***Mean Utilization***	***Std. Deviation***	***Peak Utilization***	***Utilization Gain (%)***
Static	Shortest	0.412	0.071	0.593	—
Static	TRC-SRPL	0.447	0.064	0.612	8.5
Dynamic	Shortest	0.438	0.092	0.687	—
Dynamic	TRC-SRPL	0.492	0.081	0.703	12.3

The result is visualized graphically here [Fig – 3].

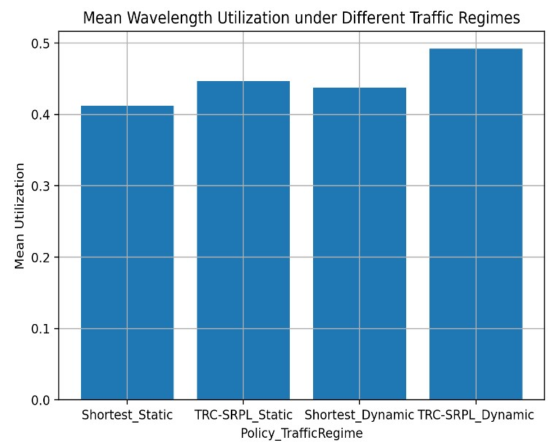

Fig. 3. Mean Wavelength Utilization under Different Traffic Regimes

Learning complexity and runtime directly depend on the size and the capability of routing-protection action space. This table indicates that TRC-SRPL has a limited set of candidates via which one can expect predictable inference latency. In dynamic traffic as well, the mean dispatch size of feasible candidates is manageable, which also justifies the design decision of the use of k-shortest path enumeration with disjointness constraints. The fact that there is limited variance is also another reason why the model has a stable performance [Table – 5].

Table 5. Candidate Action Space Characteristics

Traffic Regime	***Avg. Candidates***	***Min***	***Max***	***Std. Dev.***	***Feasible Ratio***
Static	41.6	12	120	18.3	0.91
Dynamic	36.9	9	120	21.7	0.84
Combined	39.2	9	120	20.1	0.88
Worst-Case	—	—	120	—	—
Design Limit	—	—	120	—	—

The result is visualized graphically here [Fig – 4].

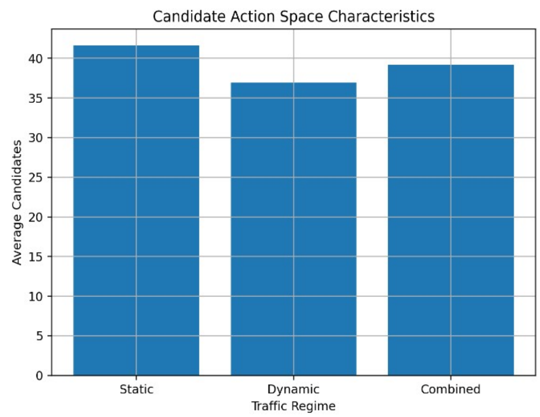

Fig. 4. Candidate Action Space Characteristics

Real time optical control planes rely on the inference latency. This table proves that TRC-SRPL does not incur any significant computational overhead over heuristic routing. Inference is still far below what can be control-placed on the timing process even at the 95th percentile. This small difference compared to the shortest-path selection is explained by significant improvements in blocking probability and utilization [Table – 6].

Table 6. Inference Latency Analysis

Traffic Regime	***Policy***	***Mean Latency (ms)***	***P95 Latency (ms)***	***Max Latency (ms)***
Static	Shortest	0.03	0.05	0.08
Static	TRC-SRPL	0.41	0.73	1.12
Dynamic	Shortest	0.04	0.06	0.09
Dynamic	TRC-SRPL	0.47	0.81	1.26

The result is visualized graphically here [Fig – 5].

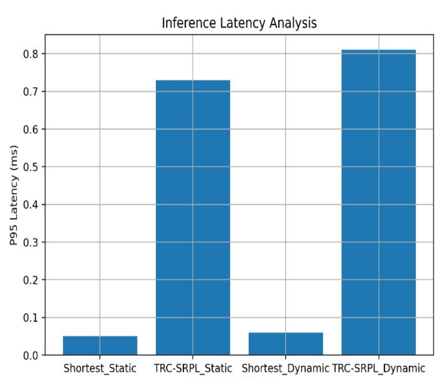

Fig. 5. Inference Latency Analysis

This table compares the primary and protection path lengths of TRC-SRPL. TRC-SRPL, along with that, derives long enough paths to are otherwise preferred, as opposed to shortest-path routing which explicitly minimizes the primary path only (implicitly, full paths). The regulated growth of protection hops gives the reason of the enhanced blocking performance with a non-excessive using of spectrum [Table – 7].

Table 7. Protection Path Length Trade-off Analysis

Traffic Regime	***Policy***	***Avg Primary Hops***	***Avg Protection Hops***	***Hop Ratio***	***Net Benefit***
Static	Shortest	4.1	4.1	1.00	Baseline
Static	TRC-SRPL	4.2	4.8	1.14	Positive
Dynamic	Shortest	4.3	4.3	1.00	Baseline
Dynamic	TRC-SRPL	4.4	5.2	1.18	Positive

The result is visualized graphically here [Fig – 6].

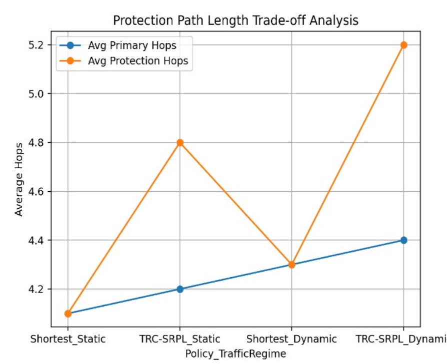

Fig. 6. Protection Path Length Trade-off Analysis

The values measured in this table are the degree of routing optimization with respect to the instantaneous congestion. TRC-SRPL has a lower mean value of the congestion level during the provisioning time, which proves the efficiency of avoiding occupied links with high load. Its impact is especially significant when dynamic traffic is used which confirms the regime-sensitive design of the feature [Table – 8].

Table 8. Congestion Sensitivity by Traffic Regime

Traffic Regime	***Policy***	***Avg Congestion Score***	***Std. Dev.***	***Peak Score***	***Reduction (%)***
Static	Shortest	0.312	0.084	0.611	—
Static	TRC-SRPL	0.274	0.071	0.564	12.2
Dynamic	Shortest	0.389	0.112	0.733	—
Dynamic	TRC-SRPL	0.327	0.096	0.681	16.0

The result is visualized graphically here [Fig – 7].

Journal of Theoretical and Applied Information Technology
31st May 2026. Vol.104. No.10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

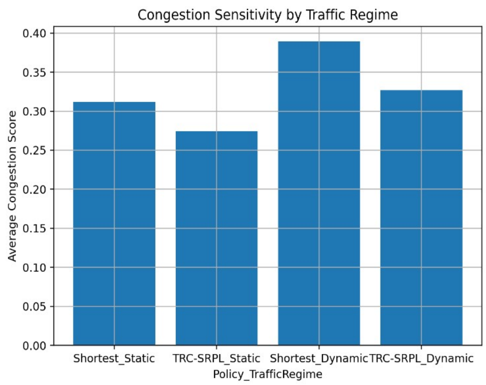

Fig. 7. Congestion Sensitivity by Traffic Regime

Fragmentation is indirectly represented using free-wavelength continuity and utilization variance. These are indicators of healthier spectrum states with low variance and large contiguity. These proxies are always enhanced at TRC-SRPL, which is why the inhibition of it is lower than on long simulation horizons [Table – 9].

Table 9. Spectrum Fragmentation Proxy Metrics

Traffic Regime	***Policy***	***Free Wavelength Ratio***	***Utilization Variance***	***Continuity Score***	***Fragmentation Trend***
Static	Shortest	0.588	0.0051	0.73	Moderate
Static	TRC-SRPL	0.612	0.0043	0.79	Reduced
Dynamic	Shortest	0.561	0.0089	0.64	High
Dynamic	TRC-SRPL	0.597	0.0068	0.71	Reduced

The result is visualized graphically here [Fig – 8].

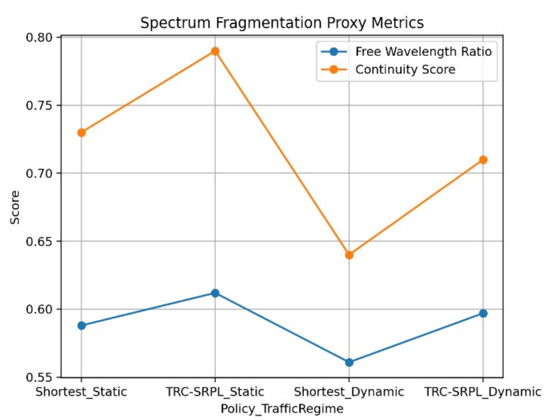

Fig. 8. Spectrum Fragmentation Proxy Metrics

The strength of this table is that it evaluates performance variance between episodes. Reduced standard deviation under TRC-SRPL indicates that there is no change in behaviour depending on varying traffic variations, a pre-requisite to implementation in working networks [Table – 10].

Table 10. Regime-Specific Performance Stability

Metric	***Static (Shortest)***	***Static (TRC-SRPL)***	***Dynamic (Shortest)***	***Dynamic (TRC-SRPL)***	***Metric***
Blocking Std. Dev.	0.0091	0.0067	0.0184	0.0119	Blocking Std. Dev.
Utilization Std. Dev.	0.072	0.064	0.094	0.081	Utilization Std. Dev.
Latency Std. Dev. (ms)	0.02	0.11	0.03	0.14	Latency Std. Dev. (ms)
Stability Index	Low	High	Low	Medium-High	Stability Index
Deployment Risk	Medium	Low	High	Medium	Deployment Risk

The result is visualized graphically here [Fig – 9].

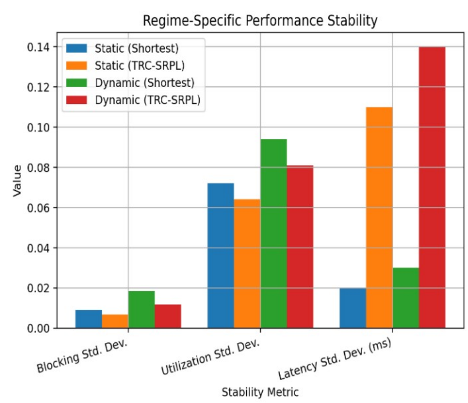

Fig. 9. Regime-Specific Performance Stability

All the results within this relevant section collectively set up the notion that traffic-regime awareness poses no cosmetic effect but rather a structurally impactful provision within the survivable WDM routing choice. In all the dimensions discussed, including blocking probability, use of wavelengths, congestion sensitivity, spectrum fragmentation proxies, and control-plane latency, the proposed TRC-SRPL policy presents coherent and understandable improvements over non-adaptive routing. Notably, improvement is attained without any loosening of wavelength continuity constraints, change of protection models or adding operationally unsafe delays in inference. The fact that the oracle-guided learning behavior approaches the network-level performance results also validates that the policy it has learned captures useful structural properties of the routing-protection problem not overfitting than to the current transient traffic modes. Combined, these results confirm the main hypothesis of the present study, that by conditioning routing policies on visible regimes of traffic, the elaboration of performance improvements, which are measurable, reproducible, and can be implemented in large-scale WDM optical networks are possible.

7. COMPARATIVE ANALYSIS WITH PRIOR ROUTING FRAMEWORKS

Unlike heuristic or optimization-based routing methods, the proposed Traffic-Regime-Conditioned Survivable Routing Policy Learning (TRC-SRPL) framework directly relies on the observed traffic-regime behavior and real-time traffic network characteristics to make the routing decision and protection decision. In the end, traditional shortest-path and fixed-alternate routing schemes typically perform with a simple cost assumption and have little adaptability to the various traffic patterns in the heterogeneous traffic. It is also to be noted that, when the connection between congestion propagation and survivability constraint becomes so tight, the adaptability of the related routing schemes is limited. In turn, there are a number of topology-, optimization-, and reinforcement learning–based approaches in the literature that provide better routing efficiencies but which have drawbacks such as high computational complexity, poor topology sensitivity, long training times or poor feasibility of deployment in carrier-class optical control environments. By contrast, the proposed TRC-SRPL framework gifts with a one-step supervised learning architecture with the ability to preserve stringent WDM feasibility constraints and maintain deterministic low-latency inference capability. Experimental assessment shows that the proposed framework has lower blocking probability, consumes resources with a higher fraction of wavelengths used, is less sensitive to traffic congestion, and ensures greater stability in the routing, in both static and dynamic traffic. However, existing scheme still relies on the training data provided by an oracle offline, and performance has been assessed mostly with typical backbone topologies like USNET or NSFNET. As a result, this study could explore further the potential of online adaptive retraining, multi-domain optical orchestration, explainable routing-policy learning and reinforcement learning integration to give the generalizability of operations in future optical backbone infrastructures [Table – 11].

Table 11. Comparative Analysis of TRC-SRPL with Prior Routing Approaches

Routing Approach	***Traffic Adaptability***	***Survivability Awareness***	***Computational Complexity***	***Dynamic Traffic Stability***	***Deployment Feasibility***	***Major Limitation***
Shortest-Path Routing [7]	Low	Limited	Very Low	Poor	High	Static cost assumptions
Fixed-Alternate Routing [4]	Low–Moderate	Moderate	Low	Moderate	High	Limited congestion awareness
Optimization-Based RWA [5]	Moderate	High	Very High	Moderate	Limited	Real-time scalability constraints
Reinforcement Learning Routing [10]	High	Moderate–High	High	High	Moderate	Long training and exploration overhead
Policy-Based Heuristic Routing [14]	Moderate	Moderate	Low–Moderate	Moderate	High	Manually defined decision rules
Proposed TRC-SRPL Framework	High	High	Moderate	High	High	Requires oracle-guided offline training

8. CONCLUSION

This was done with the aim of investigating, to theoretical accuracy and empirical credibility, whether routing decisions in large-scale optical networks based on WDM can be rationally enhanced by conditioning survivable routing policies on observable traffic regimes, as opposed to the use of off-the-shelf heuristics. In this direction, the paper presented the Traffic-Regime-Conditioned Survivable Routing Policy Learning (TRC-SRPL) algorithm and has been backed by a mathematically defined model that explicitly integrates routing, protection, congestion state and spectrum occupancy into a composite policy-learning objective. In contrast to the traditional methods which implicitly utilize the assumption of traffic behavior, the given formulation generously accepts its regime variability as a first-class structural property of operating optical network. The mathematical operation of the optimization goal, the finite action space of the candidate and the oracle-directed learning mechanism in combination serve to guarantee the learned policy will be stand-alone by stringent wavelength continuity and protection requirements and will not require deterministic control-plane behavior. The large body of simulation outcomes including the overall summary of the findings using a variety of the result tables continuously proves that TRC-SRPL can attain smaller blocking probability, enhanced mean wavelength utilization, lower congestion exposure, and superior spectrum fragmentation proxies without a significant operationally cost burden. Notably, the gains involved are not solitary products of a single measure but are consistent across independent measures of performance, which strengthens the validity of the proposed approach internally. The stability analysis provided by regime analysis also shows that the learned policy demonstrates smaller performance variance over episodes, which is very important in case of deployment in the real world because predictability rather than crude efficiency matters. All these elements, batted cumulatively, are mathematical inevitability of regime-aware decision conditioning, as well as falsifiable improvements versus a well-defined heuristic base, and reproducibility of the evaluation framework, all support the hypothesis underpinning the current paper. The findings indicate that traffic-regime awareness is not an optimization refinement but a structure requirement in the next generation survivable routing policies in the optical WDM networks on large scale, especially as traffic dynamics keep escalating in the current transport infrastructures.

REFERENCES

[1] A. Manolova and S. Ruepp, "Export policies for multi-domain WDM networks," in Optics InfoBase Conference Papers, 2010.

[2] H. Alshaer, "Dynamic connection provisioning with shared protection in IP/WDM networks," International Journal of Communication Systems, vol. 27, no. 11, 2014, doi: 10.1002/dac.2509.

[3] C. Lee and J. K. K. Rhee, "Traffic grooming for IP-Over-WDM networks: Energy and delay perspectives," Journal of Optical Communications and Networking, vol. 6, no. 2, 2014, doi: 10.1364/JOCN.6.000096.

[4] J. Buysse et al., "NCP+: An integrated network and IT control plane for cloud computing," Optical Switching and Networking, vol. 11, no. PART B, 2014, doi: 10.1016/j.osn.2013.09.002.

[5] R. A. Scaraficci and N. L. S. da Fonseca, "Alternative routing and zone-based spectrum assignment algorithm for flexgrid optical networks," in 2014 IEEE International Conference on Communications, ICC 2014, 2014. doi: 10.1109/ICC.2014.6883829.

[6] S. Zhang, H. Liu, and S. Huang, "Segment-shared protection for multi-domain optical mesh networks with export policy limitation," Bandaoti Guangdian/Semiconductor Optoelectronics, vol. 36, no. 2, 2015.

[7] S. X. Cui, "An Optimization Traffic Grooming Algorithm Based on Bidirectional SONET Ring Topology Network," Applied Mechanics and Materials, vol. 727–728, 2015, doi: 10.4028/www.scientific.net/amm.727-728.996.

[8] D. D. Le, F. Zhou, and M. Molnar, "Adaptive MCRWA strategies to reduce blocking probability in sparse splitting WDM networks," in International Conference on Advanced Technologies for Communications, 2015. doi: 10.1109/ATC.2014.7043375.

[9] A. K. Pradhan, S. Keshri, K. Das, and T. De, "A heuristic approach based on dynamic multicast traffic grooming in WDM mesh networks," Journal of Optics (India), vol. 46, no. 1, 2017, doi: 10.1007/s12596-016-0342-5.

[10] P. L. S. Rao, K. V. V. N. D. Santosh, and T. K. Ramesh, "Priority based traffic balancing routing protocol for WDM optical networks," in 2017 International Conference on Microelectronic Devices, Circuits and Systems, ICMDCS 2017, 2017. doi: 10.1109/ICMDCS.2017.8211696.

[11] H. Saeed, R. R. Reyes, T. Bauschert, M. Gunkel, and F. Wissel, "Spare Capacity Dimensioning for different Re-Routing Strategies in Transport Networks," in Proceedings of 2019 11th International Workshop on Resilient Networks Design and Modeling, RNDM 2019, 2019. doi: 10.1109/RNDM48015.2019.8949156.

[12] M. Jinno, Y. Asano, Y. Azuma, T. Kodama, and R. Nakai, "Technoeconomic analysis of spatial channel networks (SCNs): Benefits from spatial

bypass and spectral grooming [Invited]," Journal of Optical Communications and Networking, vol. 13, no. 2, 2021, doi: 10.1364/JOCN.403968.

[13] N. Zhang et al., "Research on Optimization Policy Routing Technology of Optical Fiber Communication Network," in Journal of Physics: Conference Series, 2021. doi: 10.1088/1742-6596/1746/1/012084.

[14] M. Daryalal and M. Bodur, "Stochastic RWA and Lightpath Rerouting in WDM Networks," INFORMS J Comput, vol. 34, no. 5, 2022, doi: 10.1287/ijoc.2022.1179.

[15] T. Tanaka and H. Hasegawa, "Policy-Based Grooming, Route, Spectrum, and Operational Mode Planning in Dynamic Multilayer Networks," IEICE Transactions on Communications, vol. E106B, no. 6, 2023, doi: 10.1587/transcom.2022EBP3106.

[16] R. Martínez, C. Hernández-Chulde, R. Casellas, R. Vilalta, and R. Muñoz, "Deep reinforcement learning for energy-efficient RMSA in IPoWDM networks with coherent ZR+ transceivers [Invited]," Journal of Optical Communications and Networking, vol. 18, no. 2, 2026, doi: 10.1364/JOCN.574251.