Evaluation of Deep Reinforcement Learning for Restoration in Optical Networks

C. Hernández-Chulde, R. Casellas, R. Martínez, R. Vilalta, R. Muñoz
Centre Tecnòlogic de Telecomunicacions de Catalunya (CTTC/CERCA), Castelldefels (Barcelona), Spain
chernandez@cttc.es

Abstract: A deep reinforcement learning-based agent is presented to perform autonomous lightpath restoration upon a link failure event. The agent is evaluated against other heuristic algorithms under different traffic load and failure duration scenarios.

1. Introduction

5G and B5G traffic is envisioned to be stringent and heterogeneous, requiring high capacity, low latency and ultra-high reliability. To fulfill these requirements, optical networks need to guarantee reliable data transport in an effective manner. Therefore, it is crucial that optical network control evolves towards flexible and autonomous network systems. Failure management is extremely important in optical networks since an outage of an optical connection can affect a significant number of applications or services. Similar to other control functions for optical networks, the application of Machine Learning (ML) to automate failure management functions has been a growing trend in recent years. In the literature, plenty of works explore the application of ML in optical network failure management [1]. These works address fault management from both a proactive approach which aims to prevent a network or service disruption before the failure occurs (e.g., performance monitoring and failure prediction), and a reactive approach, which acts during and after the failure intending to restore the service (e.g., failure detection, localization and identification) [1]. Nevertheless, to the best of our knowledge, the subsequent restoration process of lightpaths affected by a network failure has not been addressed through ML techniques.

This paper focuses on evaluating the feasibility of applying reinforcement learning (RL) lightpaths restoration problem when an optical link failure occurs. We propose an RL agent devised to effectively restore disrupted lightpaths. The performance of our agent is compared against a heuristic algorithm presented in [2] using as performance metrics the restorability and the blocked bandwidth ratio (BBR).

2. System Model and RL Agent for Restoration

The optical network is modeled as a graph, with the flexigrid ROADMs as vertices and C-band links as edges. Lightpath requests specifying the source, destination, and required bandwidth/optical spectrum arrive at the control system that handles resource selection and allocation. To this end, a Routing and Spectrum Assignment (RSA) algorithm is triggered to seek a feasible path meeting the lightpath requirements and fulfilling the spectrum continuity and contiguity constraints. Upon an optical link failure, all connections passing through it are disrupted. Those connections need to be efficiently and promptly restored. The lightpath restoration often involves the provision of an alternative route that satisfies the connection requirements. A set of disrupted connections is typically restored following the order reported by the network monitoring entity. This sequential approach may result in a situation where not all paths are restored even though resources are available for restoration. An alternative algorithm called Global Concurrent Restoration Routing and Spectrum Assignment (GCR) was proposed in [2]. GCR aims to seek a feasible alternative path for each disrupted lightpath, while maximizing the overall restorability. GCR shuffles the order of disrupted lighpaths to generate more candidate restoration solutions and selects the one with the highest amount of restored bandwidth. GCR performs better than the sequential approach showing an increase in restorability and a reduction of the BBR. In [3], we proposed a RL-based RSA solution that jointly considers the spectrum utilization of the optical links and the delay constraint of the service requesting the optical connection. Results demonstrated that our proposed solution outperforms the typical heuristic kSP-FF algorithm by keeping the BBR to a minimum and another RL-based solution by accelerating the learning convergence.

This work extends the RL agent we proposed in [3], to select the best restoration sequence from those supplied by the GCR algorithm. We assume that the RL agent resides in an SDN controller and has the network state information (i.e., topology and resource utilization). Upon receiving a link failure alert, the DRL agent retrieves N candidate restoration solutions from GCR. The agent's neural network (NN) is fed with a combination of the features of each candidate solution consisting of restorability (i.e., $total\_restored\_lighpaths/total\_disrupted\_lighpaths$), restorability bandwidth (i.e., $total\_restored\_bw/total\_disrupted\_bw$), and number of restored and unrestored lightpaths categorized by the required bandwidth in terms of frequency slots (FS). The output of the NN will be the selected restoration sequence. The reward received will be proportional to the restored bandwidth. This proposed RL agent is illustrated in Fig. 1 (a). It is worth recalling that RL is an iterative learning process, where the agent stores past experiences in a buffer. For each interaction with the environment, the agent stores a tuple consisting of the current state, the action taken, the reward received and the future state. The agent learns the best action from past experiences stored in the buffer, making learning a supervised learning problem.

3. Experimental Performance Evaluation

The experimental scenario supports the simulation of two main network functions: the provisioning of new connection requests and the autonomous restoration of disrupted lightpaths. Dynamic requests were generated with a random selection of end nodes for the provisioning, and their bandwidth demands were evenly distributed between [1, 2 and 4] FSs. Requests arrive following a Poisson process with mean inter-arrival time (IAT), and connection duration is exponentially distributed with a mean holding time (HT). The link failures generation is modeled as a Poisson process with a failure IAT ($FIAT$) and the failure duration follows an exponential distribution with a failure HT (FHT). The failed link is chosen uniformly among those in the 14-node NSFNET topology shown in Fig. 1 (a), where optical fiber links can accommodate 100 FSs. For provisioning, a RSA algorithm iterates over k shortest candidate paths connecting the source and destination nodes. The k paths are sorted according to the end-to-end path delay. The number of candidate paths for the RSA algorithm was set to 4 ($k = 4$).

We first analyzed the impact of failure duration with GCR under different traffic loads. To this end, FHT was set to 2, 4, 8, 16 or 32 s, and to generate traffic loads of 100, 200, 300, 400 and 500 Erlangs, HT was modified. This analysis and the further comparison were performed in terms of restorability, defined as the ratio between the number of connections that are successfully restored and the total number of disrupted connections (e.g., a restorability value of 1 means all connections are restored), and the BBR, which represents the proportion of the total requested bandwidth that cannot be provisioned (i.e., if all connections were of the same bandwidth, this is equivalent to the blocking probability). From Fig. 1 (b) and (c), it can be inferred that the higher the traffic load on the network, the lower the restoration capacity and the higher the BBR. This behavior is reasonable since a high traffic load implies more simultaneous connections occupying network resources which does complicate both the provisioning of new connection requests and the restoration of disrupted lightpaths. In Fig. 1 (b), it can be seen that there is an inverse relationship between the failure duration and the restorability. This behavior is expected since simultaneous link failures are more likely to occur with a longer failure duration. Therefore, it becomes more difficult to find available resources to establish alternative routes to restore disrupted lightpaths.

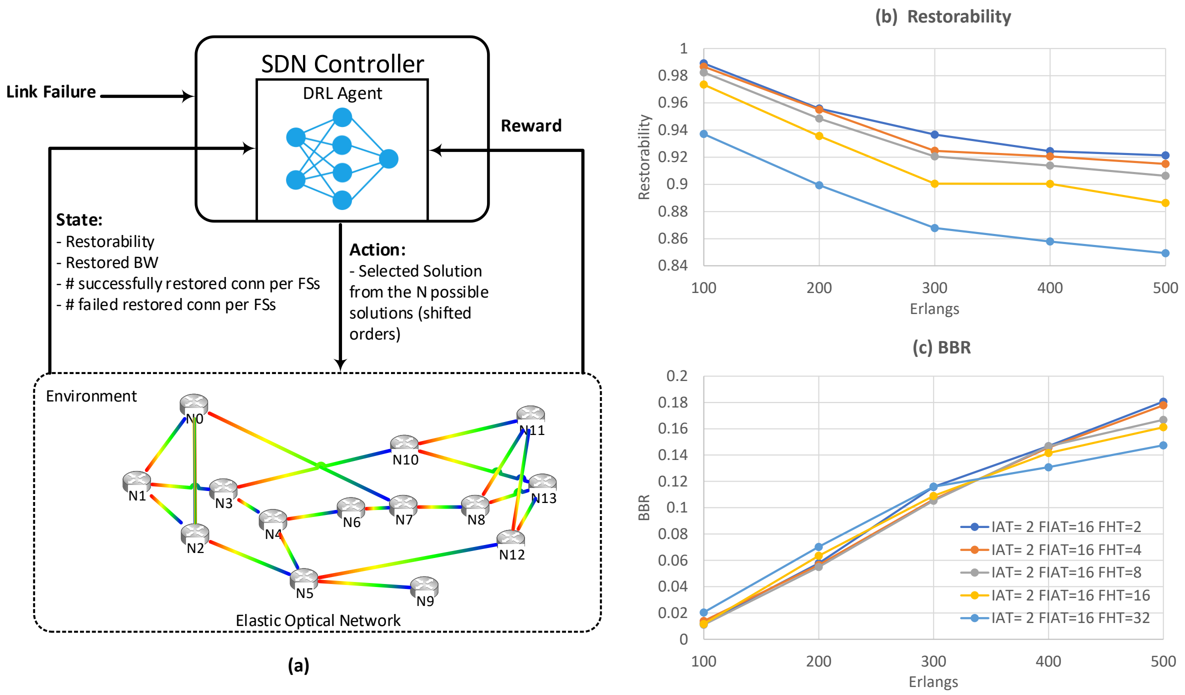

Fig. 1. (a) DRL Agent, (b) Restorability and (c) BBR under different traffic loads for GCR.

To assess the performance of the proposed RL agent, we compared it with both the GCR heuristic algorithm and the sequential approach. The DRL agent was implemented based on the Asynchronous Advantage Actor Critic (A3C) algorithm and was trained in 2000 episodes. One episode involves the provisioning of 10000 new connection requests. During provisioning, optical links fail randomly, triggering the restoration process. When a restoration runs, the agent decides which is the best restoration sequence using its neural network and saves the experience in the buffer. The size of the implemented buffer is 400. Once the buffer gets full, the agent train the neural network. The neural network consists of 5 fully connected hidden layers. After the training process, the agent deletes half of the buffer and rerun it when the buffer becomes full again.

Restorability and BBR for $IAT = 2$, $FIAT = 16$, $FHT = 2$ and 200 Erlangs are shown in Fig. 2 (a). Regarding restorability, the proposed RL-based agent improves on the sequential approach and slightly outperforms GCR. The RL agent attains an average restorability of 0.957 while GCR and sequential achieve 0.956 and 0954, respectively. Conversely, for the BBR metric, our agent only improves the performance of GCR but does not outperform the sequential approach. The BBR achieved by the agent, GCR and sequential are approximately 0.052, 0.058 and 0.048, respectively. In this case, likely, the agent does not learn accurately because most of the sample tuples stored in the buffer are very similar with a high restorability value. For this reason, the agent could not discriminate which is the best action that would improve the restorability, and thus clearly outperform the other approaches.

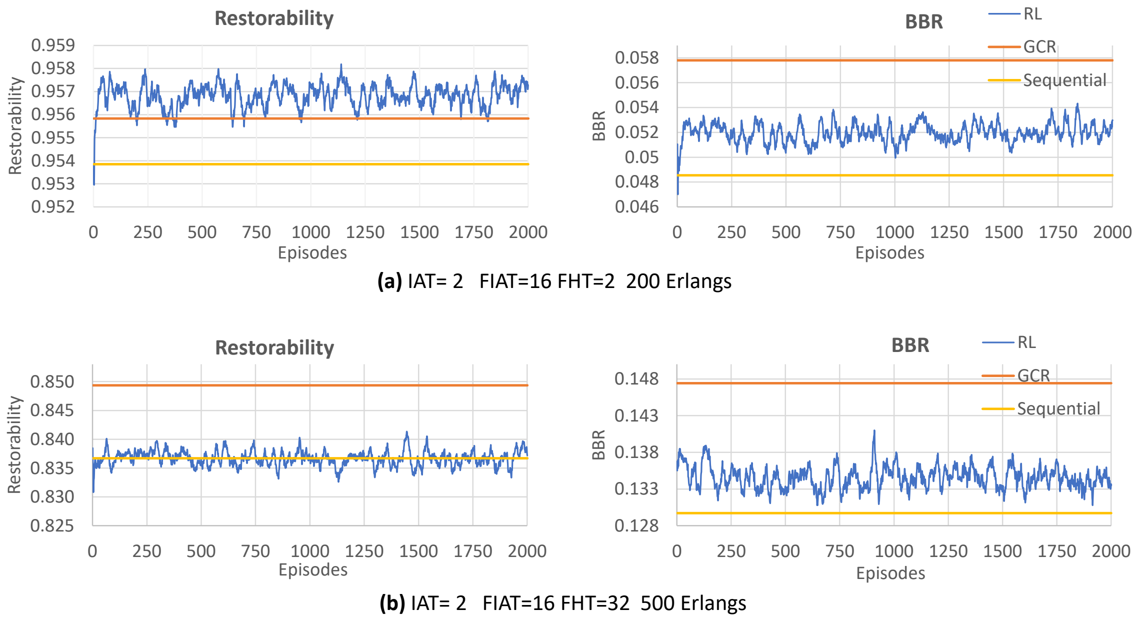

Fig. 2. Restorability and BBR using RL agent, GCR and Sequential Approach.

For $IAT = 2$, $FIAT = 16$, $FHT = 32$ and 500 Erlangs, restorability and BBR are plotted in Fig. 2 (b). The RL agent only reaches a restorability comparable to the sequential algorithm (≈ 0.837). While GCR achieves the highest restorability (≈ 0.849). In terms of BBR, the agent outperforms the GCR but does not improve the performance of the sequential algorithm. The BBR values for the agent, GCR, and sequential algorithm are around 0.135, 0.147 and 0.130, respectively. As in the previous case, the experiences stored in the buffer are very similar. Furthermore, since failure duration is longer, there will be fewer failure events, resulting in fewer samples stored in the agent's buffer. Consequently, the result is a poor performance in terms of both restorability and BBR.

4. Conclusions

This work evaluates the application of RL to restoration in optical networks. Despite the potential shown by RL in solving other networking problems, the performance of the RL-based solution at best slightly outperforms the other solutions in terms of restorability. In contrast, in terms of BBR, the RL-based solution fails to simultaneously outperform the other two approaches and, in some cases, even performs worse. This behavior is due to the fact that heuristic algorithms already offer high performance for restoration and because the system model and the number of parameters are reduced to train the agent. We expect the RL agent to perform better in scenarios involving higher complexity with a large number of variables to be considered, requiring capturing the correlation between them to infer the action to be taken.

Acknowledgements: The research leading to these results has received funding from Spanish Thematic Network Go2Edge (RED2018-102585-T), EC H2020 B5G-OPEN (101016663), MICINN AURORAS (RTI2018-099178-B-I00) and FPI scholarship PRE2019-091447.

References

[1] F. Musumeci *et al.*, "A tutorial on machine learning for failure management in optical networks," J. Light. Technol. **37**, 4125–4139 (2019).

[2] R. Martínez *et al.*, "Autonomous sdn-based global concurrent restoration for high-capacity optical metro networks," in *2021 Optical Fiber Communications Conference and Exhibition (OFC),* (2021), pp. 1–3.

[3] C. Hernández-Chulde *et al.*, "Assessment of a latency-aware routing and spectrum assignment mechanism based on deep reinforcement learning," in *2021 European Conference on Optical Communications (ECOC),* (2021), pp. 1–4.