ISSN: 2229-6948(ONLINE)
DOI: 10.21917/ijct.2022.0417

AN IMPROVISED ENSEMBLE MECHANISM FOR IMPROVING BANDWIDTH IN OPTICAL NETWORK

V. Anjana Devi[1], Rama krishna Tummala[2], E. Bhuvaneswari[3] and N. Kanagavalli[4]
[1,3,4]Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, India
[2]Accenture Solutions Pvt Ltd, Chennai, India

Abstract
An integer linear programming (ILP) model was developed with the intention of reducing the quantity of network spectrum resources while maintaining spectrum contiguity and spectrum continuity restrictions. The techniques that were just outlined have the goal of decreasing the amount of bandwidth that is used by a network; however, they do not consider the issue of spectrum fragmentation. As a result of this, many authors have zeroed in on the issue of spectrum fragmentation and presented a wide range of potential remedies. By separating a single request into many more manageable sub-requests, the Ensemble algorithm makes it less likely that a connection would be denied. However, to process the request, this will necessitate an increase in the number of guard bands and transponders that are used. This will come at a cost. EONs proposed a novel approach to the spectrum's compactness that they developed (SCS). Even though deciding on a compactness threshold can significantly cut down on the amount of spectral fragmentation that takes place, doing so is notoriously difficult for a number of different reasons.

Keywords:
Spectrum Scheme, Ensemble Algorithm, Spectrum, Bandwidth

1. INTRODUCTION

The ability to build high-capacity elastic optical networks in the future will be made possible by two key components [1]: a flex-grid ROADM that uses wavelength selective switches (WSS), which provides reconfigurable optical routing with flexible bandwidth allocation [2]-[5] and a bandwidth-variable transceiver (BVT) that is assisted by advanced digital signal processing (DSP) techniques, which accommodates varying traffic demands and link conditions. Both components are expected to be commercially available ROADM is the name given to both of these different kinds of components.

The number of cascaded ROADMs utilized for individual links in a meshed optical network might vary from one instance of that network to the next [6], the pass-band bandwidth can be subject to alterations. Because of this, the filtering penalty as well as the total link noise will be different for every single link. The capacity of standard systems to accommodate new links is extremely limited.

This is because standard systems are frequently built to withstand the most catastrophic of outcomes. BVTs can change the symbol rate and spectral efficiency (SE) of transmitted signals in several different ways, however these changes are dependent on the state of the underlying link [7]. When this is accomplished, the whole potential that is made available by each connection in a meshed optical network can be accessed and utilized.

There have been commercial implementations of flexible SE using transceivers with a fixed symbol rate and reconfigurable standard QAMs [8]. These transceivers have been used. There is now work being done to develop more advanced modulation formats that can be used in transceivers that have a finer granularity of SE. These formats include multi-dimensional modulation formats [11] as well as time-domain hybrid QAM (TDHQ) [9].

It has been demonstrated that it is possible to increase the maximum transmission distance that can be achieved for fixed data rate systems (such as 100G or 200G) by modifying the combination of symbol rate and SE utilizing BVTs in meshed optical networks with cascaded ROADM. This is the case although the data rate has not changed. It is more vital to enhance network capacity than it is to increase transmission distance; as a result, it is essential to use BVT in conjunction with a variable symbol rate and/or modulation type.

On the other side, the adaptability of BVTs comes at a cost, which manifests itself in the form of increased operational expenses as well as increased energy consumption [10]. Considering this fact, it is of the utmost importance to examine the benefits that can be delivered to the network by BVTs that have varying degrees of adaptation.

2. RECONFIGURABLE MESH SYSTEM

More than one ROADM can be used concurrently in order to filter and reroute a signal as it is traveling through a meshed optical network. As can be seen in Figure 1, the transmitting connection in such networks is portrayed as a straight line with cascaded ROADMs at various points along its length. This applies to any given pair of Tx and Rx nodes. (a). The ROADM assignment relies on the topology of the network to function properly. When ROADMs are cascaded, the resultant filtering produces a narrow pass band, which in turn limits the channel bandwidth that is available. This is because the transfer function of the ROADM is not perfectly rectangular. The reason for this can be found in the ROADM. A number of factors, such as signal bandwidth, cascaded ROADM bandwidth, and ROADM distribution in the connection, can have an effect on the amount of a toll that filtering takes on the quality of the signal. The signal bandwidth has a root raised cosine (RRC) pulse form with a roll-off ratio of 0.1, and the pulse form of the transmitted signal is also a root raised cosine. These characteristics are inversely related to the symbol rate. The filter transfer function S(f) of commercial ROADMs constructed on LCOS-based WSS can be efficiently characterized by making use of error functions.

We can calculate the 3-dB bandwidth of a ROADM given the parameters, which is independent of the other equations and does not rely on them in any way. An older 50 GHz grid, which shows how the total 3-dB bandwidth shifts as the number of cascaded ROADMs grows. In order to stop channel crosstalk from occurring, the bandwidth of a single ROADM has been limited to three decibels, or 40 gigahertz.

For BOTF values of 8.8 GHz and 14 GHz, the ideal 10-dB bandwidth of a single ROADM is 47.6 GHz, and for BOTF values of 8.8 GHz and 14 GHz, it is 51.8 GHz. The two different BOTF values that were used throughout the simulation as well as the experiment. These values were used throughout the entirety of both processes. A BOTF frequency of 8.8 GHz is used for the simulation since this frequency is more compatible with the requirements of the commercially available ROADMs [21]. As a result of the fact that their symbol rate and/or SE can be altered, BVTs are able to operate in one of four distinct modes.

This is comparable to the commercial transceiver that is already on the market. Mode 1 possesses a constant symbol rate in addition to several QAMs that are standard. For the sake of comparison, we will utilize Mode 1. The forward error correction (FEC) threshold has been set at a value of 0.02-bit error rate (BER), and the reference symbol rate is 32 Gbaud. TDHQ is used in order to provide an almost continuously varying SE while still keeping a constant symbol rate. This is accomplished using TDHQ.

The creation of TDHQ signals is accomplished by interleaving the time intervals of M1-QAM symbols and M2-QAM symbols one after the other in sequential order. Combining two normal QAM forms that are near one another results in the production of hybrid signals. Because a single TDHQ frame can contain either a single set of possible symbol indices (M1, M2) or multiple symbol indices (N1, N2), each symbol encoding rate (SE) and bit per symbol (BpS) corresponds to a different symbol. This is because a single set of possible symbol indices (M1, M2) can represent more than one symbol.

To discriminate between TDHQs and ordinary QAMs, we make use of bits per second for each polarization (BpS per polarization). Seven TDHQ signals are injected in between two conventional QAMs at a bit-per-symbol (BpS) resolution of 0.125, which allows for the simulation of a SE that is nearly always changing. This gives the idea that the SE is moving around all the time.

The difference in power ratios that are required to obtain an acceptable BER for the two hybrid forms is equal to the power ratio that exists between the two formats. This means that the difference in power ratios is equal to the power ratio that exists between the two formats. Because of the programmable symbol rate, it is possible to find a happy medium that considers both the filtering penalty and the throughput. Using TDHQ, the BVT is denoted with a variable symbol rate and a quasi-continuous SE.

This setup makes it possible to achieve maximum capacity by thoughtfully selecting an optimum combination of symbol rate and SE, considering the influence of filtering along the way. First, we perform a sweep of the symbol rate and the modulation format to locate the best combination. Next, we select the combination that results in the highest capacity that is reasonably achievable given the constraints of the system.

This enables us to determine which of the possible combinations produces the greatest results. For the purposes of this investigation, we will proceed under the assumption that all of the possible adaptation strategies have been effectively put into action. The use of flexible modulation formats and symbol rates will lead to a rise in the complexity of implementation, as well as an increase in the costs associated with power consumption.

3. PROPOSED FORMULATION

The ILP model can find a solution to the problem of shared pathways; however, it is too complicated to be implemented in big networks because of the complexity it introduces. To solve the challenge of survivable RSA with shared path protection, a new heuristic approach known as the ensemble segmentation approach has been developed. By putting into practice spectrum segmentation and reconstructing solutions that allow for shareable bandwidth, it is feasible to limit the number of spectral users and the danger of bandwidth blocking.

Spectrum resources are assigned and released by a process that is dynamic and unequal in character. This leads to the easy construction of spectrum fragmentations, which leads to a reduced usage of the spectrum, which in turn leads to a reduced use of the spectrum.

It has been hypothesized that a solution to the problem of spectrum splitting can be found through the utilization of an ensemble method. Ensemble is predicated on the concept that the spectrum may be partitioned into many segments, each of which can be utilized to communicate requests with the same priority and, as a result, bandwidth requirements.

This makes it possible to utilise the spectrum in a manner that is more effective. If the required bandwidth for the request is more than the bandwidth of the continuous spectrum block that is now available, then the request cannot be communicated inside the current spectrum segment since it exceeds the bandwidth requirements.

The protection path RSA problem can be solved using the ensemble algorithm by modifying the value of the link cost function. We decided to raise the proportion of larger spectrum blocks in the link cost function rather than just finding the cost value of each spectrum block based on the number of sharable slots, which would have been the simplest way to determine the cost value of each spectrum block. Because of this, higher precision is possible. The link cost function is defined by the standard shared path protection mechanism. This responsibility lies with the standard shared path protection mechanism.

The Reconstructing Bandwidth approach has the potential to increase spectrum consumption and lower the risk of bandwidth blocking after the network connection cost function is modified. This could be accomplished by reducing the possibility of bandwidth blocking. The conclusions of this analysis were presented earlier, and these findings are based on the findings of previous analysis. Considering that a guard band is required for each picked block, the shared spectrum blocks ought to be placed in size order, beginning with the largest and working down to the smallest, for the same reason. The order should start with the largest and work its way down.

This method that we have shown takes use of a wide number of machine learning algorithms. Some of these approaches include logistic regression, naive bayes, and random forest classifiers, amongst others. Combining the methods with a fuzzy classifier that is based on voting resulted in a significant improvement to the system's performance. This section provides an overview of the various algorithmic processes in a concise and understandable format.

3.1 LOGISTIC REGRESSION

Logistic regression is a statistical model that can be used to draw inferences between two distinct outcomes (y = 0 or 1) that are both viable options. It does this by comparing the likelihood of each of these events occurring. The approach known as logistic regression belongs to the family of linear learning methods. When using logistic regression to make predictions, it is common practice to express one's findings in terms of the probability that an event will take place. In order to assign a value to each individual data point, the LR technique makes use of the sigmoid function as an integral part of its computations.

3.2 NAIVE BAYES

This classifier makes use of statistical analysis as its technique of choice. This tactic makes use of Bayes' theorem and is predicated on the idea that the existence of one feature in a class is unrelated to the presence of another feature in the same class. Bayes' theorem was developed in the nineteenth century. To make predictions regarding the probabilities of particular categories, it is necessary to calculate the combined probability of certain categories and the words that belong to those categories.

Because of this assumption, the computations may be finished more rapidly, and it is now possible to study the parameters of each term on their own without having to rely on the others. A Bayesian network is made up of two different components, the first of which is a structural model, and the second of which is a set of conditional probabilities.

3.3 RANDOM FOREST CLASSIFIER

That is a type of classifier known as an ensemble classifier, just like the one that we have here. Improved accuracy in predictions can be achieved through the utilization of models that are based on decision trees. To extrapolate information from the training set, it creates a huge number of trees and applies the bootstrapping approach to each of those trees. This allows it to extrapolate information from the training set. During the process of classification, every tree in the forest is provided with its very own copy of the method input, and after that, each tree casts their vote independently for one of the classes. The Revision Committee (RF) will choose the class that was given the most support by the voting public to be the final class.

4. ENSEMBLE SOFT VOTING CLASSIFIER

This classifier is a meta-classifier, which means that it incorporates the predictions from multiple machine learning models of the same kind or of different sorts by voting on which predictions have most support. There are two distinct methods of voting that can be applied inside the framework of a voting classifier. Both methods are described below. The aggregater will construct the final forecast based on the class prediction that has received most votes from the base models in order to choose it. This prediction will be the one that has been used the most frequently. This result was arrived at after a round of heated debate and voting. When it comes to soft voting, it is imperative that you implement the Predict proba method into your fundamental model. This cannot be stressed enough.

As a result of the voting classifier's capacity to combine the results obtained from a variety of distinct base models, the overall performance of the system is significantly improved. The model that is now under consideration is an amalgamation of three different classifiers. The logistic regression classifier, the naive bayes classifier, and the random forest classifier are the names of these three different classifiers. A soft voting classifier has leveraged the predict proba attribute column to make its decisions. This column provides the percentage of likelihood that each of the target variables will be discovered.

In the last stage of the process, a logistic regression, Naive Bayes, or Random Forest model is utilized to conduct an analysis on the randomized data points collected from the training set. When all the predictions from the models are counted through the use of a voting aggregator and the method of soft voting, the result is the vote that received the majority of support.

5. RESULTS AND DISCUSSION

During the simulation run, there will be a total of 10^6 requests made, all of which will be processed to determine whether the methodology that has been described is accurate. The bit rates of the requests can be arbitrarily changed to one of the following values: 40 Gb/s, 100 Gb/s, 400 Gb/s, or 1000 Gb/s are the available transfer rates. It was chosen to go with the DPQPSK modulation format since it is frequently used in optical networks that have 3, 4, 7, and 16 FSs. This was the primary factor in the decision-making process. Every request has its own individual FS who oversees ensuring that it is secure. The number of requests for new services that come in is distributed according to a Poisson process, and the amount of time it takes to process these requests follows a negative exponential distribution.

The most important metric that is used in the process of ranking the algorithms is referred to as the "bandwidth blocking probability, and it is calculated by determining the ratio of the bandwidth of all blocked requests to the total bandwidth of all requests. This ratio is then used to determine the "bandwidth blocking probability. Table 1 presents the findings obtained using the metric. It has been proved that Ensemble has the lowest BBP, but it also has the lowest BBP performance. Neither of these things is ideal.

This breakthrough is the outcome of the realization that the consumption of the free spectrum can be lowered by utilizing tactics such as reconstructed shareable bandwidth and spectrum segmentation. Both strategies are credited with bringing about this insight. This, in turn, improves the possibility that protective spectrum will be shared, which in turn frees up additional spectrum resources for Ensemble.

The SU for three distinct approaches, taken under a range of different workloads, is presented in Table.1(a) to Table.1(c). The ensemble method achieves the best SU results across all possible topologies, placing it in first place ahead of the other two approaches. Take into consideration the fact that an increase in the load placed on the network results in an increase in the possibility that the FF algorithm will lead to the fragmentation of the spectrum.

In the meantime, Ensemble has a better possibility of rebuilding shareable bandwidth, which has the potential to boost SU performance. SU performance could also be improved. When it comes to ensuring that different kinds of requests are handled in the same manner, any RSA system faces a significant obstacle whenever it is obliged to do so.

The odds of each request topology being terminated at 300 Erlang, 350 Erlang, and 350 Erlang, respectively, are outlined in Table.1. When requests at 1000 Mbps are sent with 16 FSs, there is a maximum rate of blocking that can occur across all topologies. This rate is referred to as the maximum rate of blocking. Spectrum continuity and spectrum contiguity limits are the root cause of the greatest blocking values, which, in turn, reduce the likelihood of granting higher rate requests.

The method that makes use of the sharable bandwidth reconstruction methodology has the potential to improve the transmission probability for requests for higher rates. Because of the way that it functions, this advancement is feasible. For each of the three algorithms, the potential for a communication to be obstructed decreases correspondingly with the transmission speed, going from 40 Gbps to 100 Gbps to 400 Gbps, respectively. On the other hand, the algorithm can also prevent blocks from happening, albeit only to a limited amount.

Table.1(a). Performance Evaluation for paths k = 2

Algorithms	**Accuracy**	**Precision**	**F1_score**	**Recall**
Logistic Reg	76.39%	65.76%	64.08%	62.48%
K-Nearest	73.36%	59.50%	60.95%	62.48%
Support Vector	75.50%	68.58%	57.65%	49.73%
Naive Bayes	75.60%	63.14%	64.68%	66.30%
Decision Tree	72.85%	59.29%	61.44%	63.75%
Random Forest	79.03%	72.63%	65.67%	59.93%
Soft Voting	80.66%	74.59%	72.99%	71.40%
AdaBoost	76.83%	69.62%	61.33%	54.83%
Bagging	76.39%	63.75%	66.78%	70.13%
GradientBoost	76.83%	72.32%	58.93%	49.73%
XGBoost	77.27%	65.57%	67.17%	68.85%
CatBoost	76.83%	65.47%	65.88%	66.30%

Table.1(b). Performance Evaluation for paths k = 4

Algorithms	**Accuracy**	**Precision**	**Recall**	**F1_score**
Logistic Reg	99.11%	99.73%	98.94%	99.84%
Random Forest	98.57%	98.17%	99.78%	99.45%
Decision Tree	96.91%	96.45%	99.87%	98.13%
Naive Bayes	95.27%	94.04%	99.87%	96.86%
SVM	62.22%	62.22%	99.49%	77.73%
K-NN	96.36%	95.63%	99.87%	97.70%
Soft Voting	99.71%	97.73%	99.79%	99.72%
AdaBoost	97.42%	99.35%	99.63%	99.98%
Bagging Class	96.36%	99.09%	96.63%	97.85%
GradientBoost	98.02%	98.43%	99.07%	99.25%
XGBoost	81.60%	69.24%	70.93%	72.70%
CatBoost	81.13%	69.13%	69.57%	70.01%

Table.1(c). Performance Evaluation for paths k = 8

Algorithms	**Accuracy**	**Precision**	**Recall**	**F1_score**
Logistic Reg	97.65%	99.24%	97.48%	98.36%
Random Forest	97.11%	96.72%	99.29%	97.98%
Decision Tree	95.48%	95.02%	98.39%	96.68%
Naive Bayes	93.86%	92.65%	98.39%	95.43%
SVM	61.30%	61.30%	98.02%	76.58%
K-NN	94.94%	94.22%	98.39%	96.26%
Soft Voting	99.22%	96.29%	99.30%	98.25%
AdaBoost	95.98%	97.88%	99.14%	99.49%
Bagging Class	94.94%	97.63%	95.20%	96.40%
GradientBoost	96.57%	96.98%	98.59%	97.78%
XGBoost	80.39%	68.22%	69.88%	71.63%
CatBoost	79.93%	68.11%	68.54%	68.98%

6. CONCLUSION

In this study, we investigate the many dangers that face the continued existence of EONs. The issue of a low sharing degree is dealt with by the proposed Ensemble technique, which employs a defense mechanism in the form of a strategy for reconstructing shareable bandwidth to solve the problem. In addition, the approach of spectrum segmentation is helpful when it comes to the process of dividing up mass spectra. When it comes to the dynamic request provisioning in EONs, the performance of the solution that was suggested is superior because of this. In the meanwhile, the concept has been conceived of and has undergone some pilot testing; it is a fusion of EONs and SDN. Software-defined networking (SDN) makes it feasible to exercise flexible control over network traffic by decoupling the device control and data planes. This decoupling is accomplished by software. In addition to this, it offers a solid foundation for the development of new applications in addition to the further expansion of existing ones. The next step in our analysis will be figuring out ways in which the SDN concept can be enhanced in some way.

REFERENCES

[1] X. Zhou and D.V. Plant, “On the Capacity Improvement Achieved by Bandwidth-Variable Transceivers in Meshed Optical Networks with Cascaded ROADMs”, *Optics Express*, Vol. 25, No. 5, pp. 4773-4782, 2017.

[2] H.S. Abbas and M.A. Gregory, “The Next Generation of Passive Optical Networks: A Review”, *Journal of Network and Computer Applications*, Vol. 67, pp. 53-74, 2016.

[3] V. Kumar and E.J. Schwabe, “Improved Access to Optical Bandwidth in Trees”, *Computer Networks*, Vol. 97, pp. 437-444, 1997.

[4] M. Rajalakshmi, V. Saravanan and C. Karthik, “Machine Learning for Modeling and Control of Industrial Clarifier Process”, *Intelligent Automation and Soft Computing*, Vol. 32, No. 1, pp. 339-359, 2022.

[5] Y.H. Robinson, V. Saravanan and P.E. Darney, "Enhanced Energy Proficient Encoding Algorithm for Reducing Medium Time in Wireless Networks", *Wireless Personal Communications*, Vol. 119, No. 4, pp. 3569-3588, 2021.
[6] X. Zhou and D.V. Plant, "Bandwidth Variable Transceivers with Artificial Neural Network-Aided Provisioning and Capacity Improvement Capabilities in Meshed Optical Networks with Cascaded ROADM Filtering", *Optics Communications,* Vol. 409, pp. 23-33, 2018.
[7] T. Thangappan, A. Suvarnamma and G.S. Swapna, "Review on Dynamic Bandwidth Allocation of GPON and EPON", *Journal of Electronic Science and Technology*, Vol. 18, No. 4, pp. 100044-100054, 2020.
[8] X. Bai, A. Shami and C. Assi, "On the Fairness of Dynamic Bandwidth Allocation Schemes in Ethernet Passive Optical Networks", *Computer Communications*, Vol. 29, No. 11, pp. 2123-2135, 2006.
[9] M. Matinmikko, P. Ahokangas and M. Mustonen, "Bandwidth Sharing using Licensed Shared Access: The Concept and its Workflow for LTE-Advanced Networks", *IEEE Wireless Communications*, Vol. 21, No. 2, pp. 72-79, 2014.
[10] E. Hossain, D. Niyato, and Z. Han, "*Dynamic Bandwidth Access in Cognitive Radio Networks*", Cambridge University Press, 2009.
[11] N. Tasneem and M. Hossen, "Traffic Localization-Based Dynamic Bandwidth Allocation Algorithm for Improvement of QoS in Peer-to-Peer Architecture of Passive Optical Networks", *Optical Fiber Technology,* Vol. 67, pp. 102722-102743, 2021.