# OBS network blocking probability prediction using ensemble technique

Srija Chakraborty
*Dept. of C.S.E.*
*National Institute of Technology*
Rourkela, India
518CS6016@nitrkl.ac.in

Ashok Kumar Turuk
*Dept. of C.S.E.*
*National Institute of Technology*
Rourkela, India
akturuk@nitrkl.ac.in

Bibhudatta Sahoo
*Dept. of C.S.E.*
*National Institute of Technology*
Rourkela, India
bdsahu@nitrkl.ac.in

2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC) | 978-1-7281-8880-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/iSSSC50941.2020.9358862

***Abstract*—In present time, artificial intelligence techniques are widely used in network field to improve the performance and security of optical burst switching networks. In our work we have used different predictor models and feature reduction algorithms to predict the blocking probability of upcoming traffic in a OBS network which will help us improve the performance of the network drastically. In this work, ensemble of predictors like Linear Support Vector Machine (SVM), Linear Regression (LR) and Random Forest (RF) are used to train and predict burst contention or burst blocking probability. From the result, we can conclude that this approach works better and gives accurate as well as faster result than other approaches. We observe the proposed model competently and predict the blocking probability with higher accuracy, along with lowering the number of burst loss. Thus, it will help future network designers to have an advanced idea about the performance of the network under specific framework.**

***Index Terms*—Optical burst switching networks, Blocking probability, Ensemble, Linear support vector machine, Linear regression, Random forest**

## I. Introduction

Optical burst switching network [1] provides buffer-less transmission and large bandwidth for data transmission, which ideally fulfill the present day's communication requirements. We use this switching technique for a bigger network with higher traffic. This switching technique performs better than the other two existing switching techniques as it overcomes the disadvantages of OPS and OCS. By using OBS, we can avoid the extensive amount of RAM requirement problem of OPS. Also for OCS, high wavelength consumption issue can be resolved in OBS network. We also use OBS for secure data transmission. In present days, SDM (Space division multiplexing) is proposed for OBS network. However in our work, we have used wavelength division multiplexing (WDM) [2] which is not time slotted in nature as it gives higher network utilization and lower the blocking probability than SDM multiplexing. In OBS [3], the time difference between the burst control packet (BCP) transmission and data burst transmission is called offset time. If transmission time of BCP is greater than the offset time, then there will be a possibility of burst contention and burst drop. Burst contention occurs whenever at the time of transmission burst gets blocked. This is the major disadvantage of OBS network. Therefore, if the blocking probability can be predicted for a network, then the performance of the network can be improved. The blocking probability of any network can be determined by calculating the proportion of the count of the bursts (lost) divided by the aggregate count of burst (sent). For OBS, network blocking can be categorised into two types. One is receiver blocking and another is channel blocking. In this work we have used wavelength conversion [4] to handle receiver blocking. In wavelength conversion, output bursts are transmitted to other output ports when the allocated output port is not free. It is much more flexible than other existing strategies.

Using a different analytical approach and simulation, we can calculate blocking probability for a network. To calculate the blocking probability simulation, we can use any event-driven simulator and for analytical approach, Erlang fixed point approximation (EFPA) [5] is used. However, these traditional approaches are much more time-consuming and tedious. Therefore, different artificial intelligence techniques are used for prediction.

We have used standard predictors like linear support vector machines, linear regression, random forest and apply bagging and predicting final result using voting technique. In this work, we have also compared our result with traditional approaches and from the results, we can state that our model has much less error than other existing traditional approaches even in higher traffic load. We have created the dataset after analysing the OBS network. Then for reduction of features in our dataset, we have used wilcoxon signed-ranked test and principal component analysis procedures. Both feature reduction techniques are used separately to create separate datasets to observe that, which technique performs better with our proposed ensemble model. Both the feature reduced datasets are used to perform the model training separately. The datasets perform differently in the training model.

The major contributions of our paper are:

1) To predict the blocking, we have created a new dataset by analysing our implemented OBS network with wavelength conversion strategy in OMNET++ version 4.3.
2) On our dataset, we have considered different features of the network. However for faster prediction, we have used feature reduction techniques like wilcoxon signed-rank test and principal component analysis (PCA) to remove least important features.
3) We have gained 97% accuracy using principal component analysis and 80% accuracy using wilcoxon signed-rank test at the time of prediction of blocking probability.

We organise the remaining part of this paper as follows. Section II briefly reviews related work. Section III describes our proposed technique. Section IV describes the results from

This work was supported by SERB, DST project.

978-1-7281-8880-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.

the experiments and discusses the comparison with related work. Finally, conclusions and future work are stated in Section V.

## II. Related Work

In this section we have discussed the work done till now.

### A. Optical Burst Switching Network(OBS)

In optical burst switching network [6], burst is formed and transmitted. Among other existing switching techniques, OBS [7] performs better for network with high traffic. Out-of-band signalling makes OBS channels more secure to transmit data. Control burst packet is transformed using control channel and for data bursts, data channel is used. Before transmitting data burst, it transmits a burst prediction control packet to inform the intermediate nodes about the arrival of the data burst. An offset time is maintained between burst control packet transmission and data burst transmission. Different machine learning techniques [8] are used to solve disadvantages of optical burst switching network.

### B. Wavelength Conversion Strategy

There are many burst contention resolution strategies [9], [10] present in optical burst switching network like wavelength conversion, Fiber Delay Lines (FDL), deflection routing, burst segmentation. In our work, we have considered wavelength conversion for our network. As this is a very flexible and less complicated strategy. For better result in our work, we have considered that all bursts are allocated with same priority. In this burst contention resolution strategy, data burst is diverted to another output port if the allocated output port is already occupied.

### C. Neural Network

In networking, neural network [11] models are used to solve issues efficiently. Multi-layer perceptron model [12] works efficiently in the matter of evaluation speed. Therefore, to predict blocking probability, neural network models can be used.

### D. Linear Support Vector Machine

The Support Vector Machine (SVM) is discriminatory in nature and formally defined by a separating hyperplane. In this, $w^T x$ is the hyperplane of linear SVMs [13] which segregate data points of two classes optimally. Where $w$ denotes the hyperplane that learns from the training data using stochastic gradient descent (SGD) method. SGD is used to optimize datasets with higher dimension [14]. In this algorithm, the objective function used in SGD as a feature vector, $x_i \in X$ and respectively $y_i$ lies between 0 to 1. A regularisation constant $\alpha$ is used to penalise the model with a higher complexity. Therefore, the loss function $L$ is used to determine the objectives of SGD. For soft margin SVMs, $L$ is denoted as higher loss and is represented as,

$$L(t, y) = max(0, 1 - t_y) \tag{1}$$

Also, the objective function will be,

$$E(w) = \frac{1}{p} \sum_{i=1}^{p} L(y_i, w^T x_i) + a||w||_2 \tag{2}$$

During this prediction, a sample is labelled as "negative" when,

$$w^T > \Lambda \tag{3}$$

### E. Linear Regression

Linear regression model [15], [16] and [17] is preferable as prediction results are the probabilistic estimates. This equation of this model is represented using Equation 4.

$$Y = aX + b \tag{4}$$

where $X$ is the independent variable, $b$ is the slope and $a$ is the y-intercept. The value of $a$ can be calculated using Equation 5.

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \tag{5}$$

The value of $b$ can be calculated using Equation 6.

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \tag{6}$$

### F. Random Forest

When a decision tree is trained using bagging or pasting method, it is called random forest [18], [19] and [20]. In this model, at the time of the training, the training set is set to max_samples. Using random forest algorithm in each step, when the tree grows, extra randomness is introduced. In this algorithm, best features are not searched while splitting a node rather than it is searched among a random subset of features. Because of this quality, we get greater tree diversity by trading off higher bias for low variance. However, this model works overall well. The classification formula for this model is:

$$J(k, t_k) = \frac{m_{left}}{m} G_{left} + \frac{m_r ight}{m} G_{right} \tag{7}$$

Where $G$ evaluates the impurity of the subsets left and right and $m$ represents the total number of instances available in the subsets.

### G. Bagging

In bagging, we use same predictor many times and the training instances are sampled. After it completes the training of the predictor, aggregating the predictor instances prediction is made. For classification of the statistical model, aggregation function is used. Bias of each predictor is higher than the model trained on the actual training set, but both variance and bias can be reduced using aggregation. Therefore, in the result, we get ensemble model with similar bias but small variance than a single predictor trained on the original training set.

### H. Voting

After using each model, we will get accuracy for each of them. Using each of these results, we can make a better prediction by aggregating the predicted results and by voting for the final class, which has more number of votes. This is a much easier way. When a test case is made considering majority votes, we call it hard voting. This gives higher accuracy than the best result of any model in the ensemble. If ensemble consists of different weak and diverse models, then also this plays the role of strong learner.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.

### *I. Wilcoxon Signed-Rank Test*

We can compare samples which are paired or related using this non parametric statistical procedure. The null hypothesis ($\mu_d = 0$, where $\mu_d$ is the location parameter of distribution of differences) in the Wilcoxon signed-rank test [21] is that a probability distribution of the set of pair-wise differences is centered at zero. For this procedure, differences comes from a continuous distribution while symmetric distribution is one of the key assumption here. Alternative hypothesis also exists here, where $\mu_d > 0$. Test statistics for this procedure are calculated by either adding the ranks assigned to the positive differences ($T_+$) or by adding the ranks assigned to the negative differences ($T_-$). If there are n differences, then the two outcomes are related as:

$$T_- = \{\frac{[n(n+1)]}{2}\} - T_+ \tag{8}$$

For analysis of the null hypothesis, a rejection region can be formed for the test statistic, $T_+$. This region can be formed from the exact null hypothesis distribution of $T_+$. With the help of permutational argument, null distribution can easily be determined for each of the possible configuration of signs $(+or-)$.

### *J. Principal Component Analysis*

Principal component analysis [22] is used to convert a set of observations of possibly co-related variables to a set of values of linearly uncorrelated variables using an orthogonal transformation. The primary goal of this technique is to decrease the dimensionality of data maintaining variance of the original dataset. In this technique, pattern is identified and data are categorised based on their similarities as well as differences using eigen values and eigen vectors. In this technique, following equation is used to achieve the decreased dimensional data,

$$Y = Z^T X \tag{9}$$

Where X is the original dataset and Z is the orthogonal matrix containing eigen vectors.

## III. Proposed Technique

In this section, we propose an ensemble network model using predictors like SVM, linear regression, random forest, bagging and voting for predicting the blocking probability of OBS networks. We present the detailed architecture of our proposed model in Figure 1. It involves three steps: creation of a dataset of OBS network by running simulations, feature reduction using wilcoxon signed-ranked test and principal component analysis separately and prediction from the datasets using the proposed ensemble model. The proposed technique is represented in Algorithm 1. The above three described steps are in-depth discussed given below:

### *A. Dataset creation*

In this work, A ring OBS network with ten nodes( node $1, 3, 5, 7, 9$ is considered as edge node and node $2, 4, 6, 8, 10$ is considered as core node) is considered, as shown in Figure 2. For our network, we have considered seven independent parameters as input, shown in Table I: E, D, C, W, $T_{off}$, APL and CR. We have considered only these seven parameters, as the number of the parameter is directly proportional with the training time of a network. A huge training time is not workable in actual time. These input parameters work independently at the time of simulation. The parameters are defined as: E is considered as the major factor to define the blocking probability, as it is the mean traffic load per source to destination (S-D) pair. Traffic load, which is measured in Erlang is directly proportional with the blocking probability. D is the traffic load difference between maximum and minimum traffic load in source to destination pair. With this parameter, traffic load variation can be understood in the network. C is denoted by a number of channels for every wavelength.

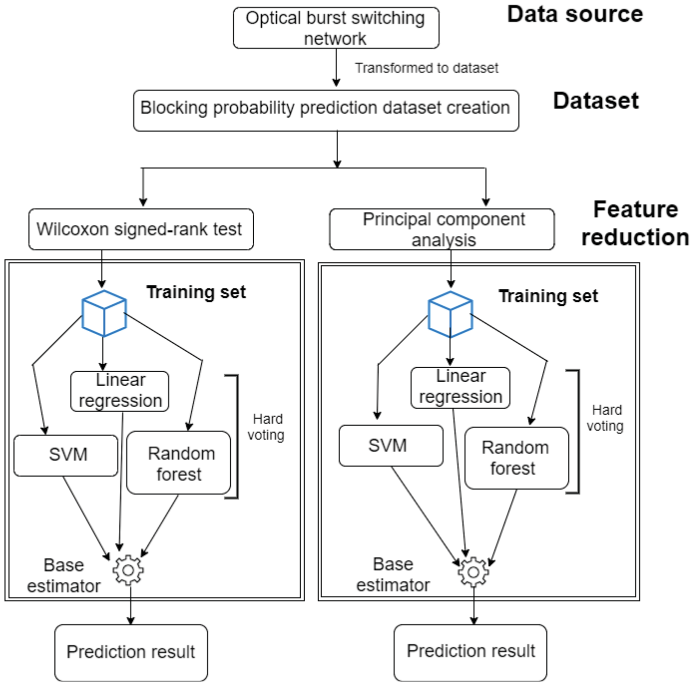


Fig. 1. Proposed architecture of our approach

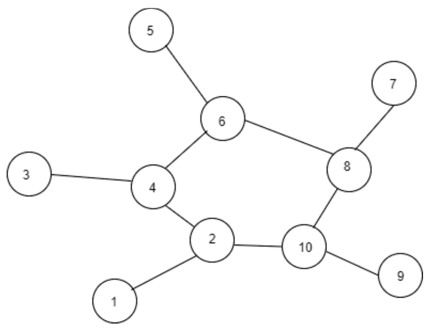


Fig. 2. Ten-node ring topology network

TABLE I
INPUT PARAMETERS IN THE SIMULATION

| Parameter | Definition |
|---|---|
| E | Mean traffic load in S-D pair |
| D | Difference between max & min traffic load |
| C | Number of channels per wavelength |
| W | Number of wavelengths |
| $T_{off}$ | Offset time |
| APL | Average path length |
| CR | Concentration of route |

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.

When F fibers are present and W wavelengths are present in each fiber, then the S sub wavelength of each fiber C will be calculated as $F * S$. Offset time $T_{off}$ is defined as the time gap between data burst and burst control packet. Whenever the transmission time of burst control packet is greater than the offset time, there will be a possibility of burst drop and blocking. We define APL as the mean distance, average number of hops and average traffic between source to destination pair. In the OBS network, CR is denoted as load balancing indicator. It is used to calculate the average number of routes. To estimate the blocking probability, parameter E is mainly used. With the help of the above described input parameters, our proposed ensemble model of the network is trained to study the mapping from the condition of the network to predict blocking probability.

**Algorithm 1:** Ensemble blocking probability prediction model for OBS network

```
Input: A data set X
Output: A predictor model which predicts blocking
        probability of OBS network.
Reduce unwanted features using wilcoxon signed-rank
  test X_W;
Reduce unimportant features using PCA X_P;
Function Training(X_W, X_P):
    for training instances in dataset X_W do
        Apply bagging on SVM, LR, RF one by one
          on dataset X_W;
        Apply hard voting to predict final blocking
          probability W_B;
        W_B ← blocking probability decided using hard
          voting ;
    for training instances in dataset X_P do
        Apply bagging on SVM, LR and RF one by
          one on dataset X_P;
        Apply hard voting to predict final blocking
          probability P_B;
        P_B ← blocking probability decided using hard
          voting ;
    return;
end function
```

### B. Feature reduction

In this work, we have considered two different feature reduction techniques for a better result. Both techniques are used in our work in pipeline.

*1) Wilcoxon Signed-Rank Test :* This technique is used for the analysis of matched pair data. After creating different sample sets, this technique tests the hypothesis of matching probability distribution of different data sets. For large samples, Z is used as standard normal distribution to test the hypothesis. Here, a two-tailed rejection region for the null hypothesis based on T+ is used:

$$Z_{+} = \frac{\frac{[T_{+} - n(n+1)]}{4}}{\{\frac{[n(n+1)(2n+1)}{24}\}^{\frac{1}{2}}} > Z_{1-\frac{\alpha}{2}} \tag{10}$$

TABLE II
NUMBER OF FEATURES USED IN EVERY DATASET CREATED

| Name | Number of features used |
|---|---|
| Original dataset | 19 |
| Wilcoxon signed-rank test | 18 |
| Principal component analysis | 15 |

or,

$$Z_{-} = \frac{\frac{[T_{-} - n(n+1)]}{4}}{\{\frac{[n(n+1)(2n+1)}{24}\}^{\frac{1}{2}}} > Z_{1-\frac{\alpha}{2}} \tag{11}$$

To prove the hypothesis, a one-tailed test will also be conducted in a similar fashion with the comparison made to $Z_{1-\alpha}$.

*2) Principal Component Analysis:* PCA is used to decrease the dimension of the data by assessing a couple of orthogonal linear combinations (LCs) of the initial variables with the largest variance. For first LC, $y_1$ is the linear combination with the largest variance. Also, $y_1 = x^T v_1$, where $v_1$ is the coefficient vector with dimension n

$$v_1 = argmax_{||v=1||} Var\{x^T v\} \tag{12}$$

Where, Var is the variance of the vector. Like that, the second LC is the linear combination with the second largest variance. Number of LCs and number of variables are same in the given dataset. The most of the LCs of datasets show variance, so that the remaining can be ignored with negligible loss of information. The scale of the variables are responsible for the variance, first standardization of each variable having mean zero and standard deviation one. When the standardization is done, the distinctive estimation units with original variables are all in equivalent units. Therefore, the experimental covariance matrix will be,

$$\sum = \frac{1}{m} X X^T \tag{13}$$

With the help of spectral decomposition theorem,

$$\sum = Z \Lambda Z^T \tag{14}$$

Where $\Lambda$ is the matrix which is diagonal in nature containing ordered eigen values. Using these, finally the output can be calculated as,

$$Y = Z^T X \tag{15}$$

*3) Modified Dataset:* Both this above mentioned feature reduction techniques are used in pipeline on the dataset obtained from OBS network. Total 19 features were present in our original dataset. After preforming the reduction processes, feature number was decreased by removing unwanted features. After using wilcoxon signed-rank test, we get 18 feature in the modified dataset. After using principal component analysis, 4 unwanted features was removed in the modified dataset. Therefore, from Table II, we can conclude that for this dataset of OBS network, PCA reduced more unwanted features than wilcoxon signed-rank test.

### C. Training Model

After feature reduction, we get two different modified data sets from two different techniques. Both the modified datasets are then trained using ensemble training models separately. Each of the model train the dataset separately and then using base estimator, prediction result is obtained. The models, we

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.

have used are linear SVM, linear regression, and random forest. We have used each of these models separately as the number of instances in a network is higher, so we prefer to use random forest since it works efficiently on large datasets. We are using the ensemble technique to predict the final burst probability, in order to do that, a weak learner is required which is SVM and linear regression. We train the model utilizing linear regression. With linear regression, the samples can be predicted with higher or lower accuracy. We have applied the hard voting on them, which have been used in our work to train our datasets.

From Figure 1, we can observe that the proposed model uses two techniques i.e. bagging and hard voting. To acomplish ensemble, we use bagging and as we predict the exact result instead of probable result, we use hard voting. In our model, hybridization is done by bagging. After training is performed utilizing the linear regression, next we train the model using a linear support vector machine. A Support Vector Machine (SVM) is a very strong and versatile machine learning model capable to perform linear as well as nonlinear classification. Linear SVMs work efficiently most of the times, many datasets are not even being linearly separable. For getting that situation to regularise the SVM, we have set C=1 and loss hyper-parameter to “hinge”. At the last stage of algorithm during designing the base estimator, random forest algorithm is used. The parameters which are used in random forest are the maximum depth of tree, random state and also the minimum number of samples needed for splitting an internal node and the minimum number of samples needed to be at a leaf node. In the case, where maximum depth of the tree is set to the default value i.e. 0, then until the leaves are pure or contain less than the minimum number of samples required to split an internal node, nodes keep expanding. In our model, the tree’s depth is set to 10. To generate a random number, random state instance is used. In our model, which has been kept at 42. For building a tree, we have set a minimum value of 1000 samples. All the selected samples should be considered at the time of splitting the internal node. In a tree, if it leaves at least minimum training samples in each of the left and right branches, then it is considered as a split point at any depth. For less than 500 samples for left or right branch, splitting can not done. In our work, sample dataset of 10,000 instances is considered and we are using our proposed ensemble algorithm. Bagging of the proposed algorithm works as described below.

1) Create random sub-samples of our dataset with replacement.
2) Train our proposed model on each sample.
3) For every new dataset, from each model calculate the average prediction.

In our proposed architecture, the ensemble model used for training each sample using a combination of linear regression, SVM and Random forest. After that, the sampled training set results are used to the base estimator obtained from hard voting phase. Hard voting derived the best result by selecting the best result calculated using individual algorithms. This considers the sampled training set as input to make predictions for them. When bagging is accomplished successfully, we inspect the performance of our proposed model to check final blocking probability calculation.

## IV. Experimental Results

The proposed ensemble model is trained and used to estimate the blocking probability of a OBS network. In this ensemble model, linear regressor, linear SVM and random forest is used. The dataset, we have used for our work has a wide range. While applying ensemble with wilcoxon signed-rank test technique as feature reduction, the root-mean-square error($RMSE$) is calculated as $7.4 \times 10^{-4}$. Similarly, with PCA technique as feature reduction, the root-mean-square error($RMSE$) is calculated as $2.3 \times 10^{-4}$. We have successfully plotted blocking probability for different values of our parameters $D$, $C$, $W$, $T_{off}$, $APL$ and $CR$. Also, the result of our model is compared with the simulation result obtained from OMNET++ simulator and result obtained using analytical approach (EFPA). We have considered two different well-known feature reduction techniques to reduce unwanted features from the dataset. From these techniques, two different datasets are obtained. In this work, we have observed performance of both the datasets using the ensemble proposed model. The first feature reduction technique used is wilcoxon signed-rank test which does not perform well for this model and gives only 80% accuracy. However, for the second technique which is principal component analysis, it performs much better and gives 97% accuracy for our dataset obtained from OBS network.

In Figure 3, we have considered $D$ = 6, $C$= 40, $W$= 3, $T_{off}$= 0.0000086, $APL$= 3.75, $CR$= 3. From the graph, we can observe that ensemble model does not work well with wilcoxon singed-rank test. However, it works extremely well with PCA and performs better than EFPA and gives almost accurate result. For lower traffic load, proposed ensemble model with wilcoxon singed-rank test performs better than the other two models. But for higher traffic load, proposed ensemble model with PCA works much better than both EPFA and proposed ensemble model with wilcoxon signed-rank test model and gives almost accurate result. For EPFA, for lower traffic load it perform almost like proposed ensemble model with PCA. However, for moderate traffic it underestimates the result and for higher traffic it overestimates the result.

In Figure 4 values of the parameters are considered as, $D$ = 8, $C$= 50, $W$= 2, $T_{off}$= 0.0000145, $APL$= 2.628, $CR$= 2.25. From the Figure 3 and 4, we can observe that for lower traffic load all the models perform poorly and gives separate results. However, when the traffic load is increased, proposed ensemble model with PCA performs better and gives much more accurate result than other two models. Therefore, considering both the results and graphs we can conclude that proposed ensemble model does not works well with wilcoxon signed-rank test feature reduction technique. However, it performs extremely well with PCA feature reduction technique. For EFPA, when the traffic load is moderate in nature then it performs well. However, for lower or higher traffic load it does not perform accurately.

From both the figure, we can conclude that when the network is booting up, the variation of blocking probability will be very high. Because of that reason, our model can not perform well. However, when the traffic load increases, the network becomes stable and fluctuation of burst loss will be in defined limit. In this range, our model performs best.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.

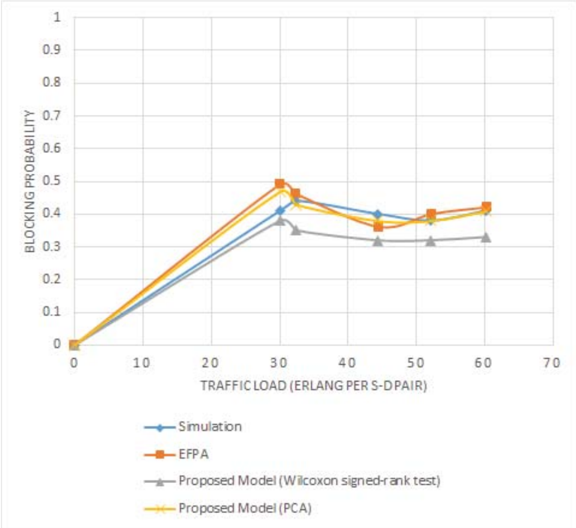


Fig. 3. Blocking probability estimation by Simulation, EFPA, proposed ensemble model with wilcoxon signed-rank test and proposed ensemble model with PCA with setting D = 6, C= 40, W= 3, $T_{off}$= 0.0000086, APL= 3.75, CR= 3

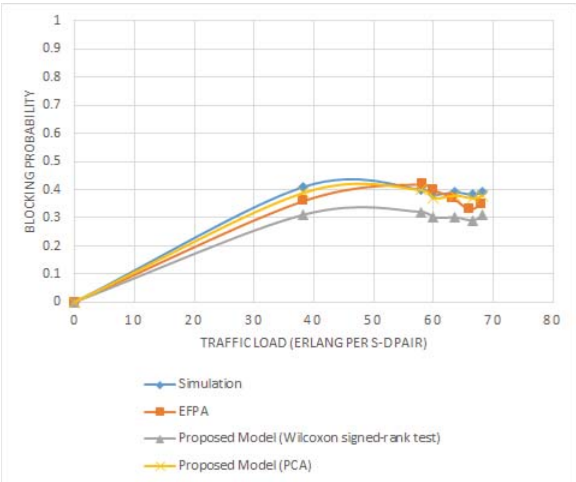


Fig. 4. Blocking probability estimation by ensemble proposed model, ADMM-I-ELM, ADMM-Log-I-ELM, EFPA, Simulation with setting D=10, C=18, W=5, $T_{off}$= 0.000009, APL=2.628, CR=2.25

## V. Conclusions and Future work

In this work, we have presented an ensemble model to train and estimate the blocking probability of optical burst switching network. For feature reduction, we have used both wilcoxon signed-rank test and principal component analysis. For training, linear support vector machine, random forest as well as linear regression is used. We have compared our work with the analytical approach of the solution and simulation results for OBS network using OMNET++ simulator. Upon observing the results, we can conclude that the model with PCA, which is used in this work gives a better, accurate and consistent result than other methods which are considered for comparison. Though for the data set, where wilcoxon signed-rank test is used, does not perform well. In this work, wavelength conversion is used as it is the most flexible approach of burst contention resolution strategy for OBS network. Estimation by EFPA method also drastically changed with the change of parameters.

In the future, our proposed concept can be implemented for other networks as well. Ten nodes are used to build our proposed network. Number of nodes can be increased for the network. In this network, we have used ring topology, other topologies can be used for implementation and experiment for this proposed model. In the future, using OBS network, other neural network approaches and model can be implemented and compared with our proposed ensemble model.

## Acknowledgment

This work was supported by SERB, DST project, and SERB grant EMR/2016/005448.

## References

[1] J. H. Ortiz, *Telecommunications Networks- Current Status and Future Trends*. InTech, 2012.
[2] Y. Liu, H. Yuan, A. Peters, and G. S. Zervas, "Comparison of sdm and wdm on direct and indirect optical data center networks," 2016.
[3] M. Nandi, A. K. Turuk, B. Sahoo, D. K. Puthal, and S. Dutta, "A new contention avoidance scheme in optical burst switch network," in *2008 16th IEEE International Conference on Networks*. IEEE, 2008, pp. 1–5.
[4] R. Lao and R. Killey, "Design of wavelength-routed optical network topologies to minimise lightpath blocking probabilities," in *Proc. London Communications Symp.*, 2003.
[5] P. Garg and N. Goyal, "Congestion control using modified erlangs loss formulae," *Int. J. Eng. Res. Technol.*, vol. 1, no. 6, pp. 1–4, 2012.
[6] Y. Xin and L. Battestilli, "A performance study on optical burst switched networks: the ring topology," *Photonic Network Communications*, vol. 14, no. 1, pp. 63–70, 2007.
[7] R. S. Barpanda, A. K. Turuk, and B. Sahoo, "Qos aware routing and wavelength allocation in optical burst switching networks using differential evolution optimization," *Digital Communications and Networks*, vol. 4, no. 1, pp. 3–12, 2018.
[8] R. Gu, Z. Yang, and Y. Ji, "Machine learning for intelligent optical networks: A comprehensive survey," *Journal of Network and Computer Applications*, vol. 157, p. 102576, 2020.
[9] R. Lamba and A. K. Garg, "Survey on contention resolution techniques for optical burst switching networks," *International Journal of Engineering Research and Applications*, vol. 2, no. 1, pp. 956–961, 2012.
[10] R. Jankuniene and P. Tervydis, "The contention resolution in obs network," vol. 20, no. 6, 2014.
[11] H. C. Cho, M. S. Fadali, and H. Lee, "Neural network control for tcp network congestion," in *Proceedings of the 2005, American Control Conference, 2005*. IEEE, 2005, pp. 3480–3485.
[12] D. R. de Araújo, C. J. Bastos-Filho, and J. F. Martins-Filho, "Methodology to obtain a fast and accurate estimator for blocking probability of optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 5, pp. 380–391, 2015.
[13] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.
[14] T. Zhang, "Solving large scale linear prediction problems using stochastic gradient descent algorithms," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 116.
[15] A. Le Calvé and J. Savoy, "Database merging strategy based on logistic regression," *Information Processing & Management*, vol. 36, no. 3, pp. 341–359, 2000.
[16] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013, vol. 398.
[17] G. King and L. Zeng, "Logistic regression in rare events data," *Political analysis*, vol. 9, no. 2, pp. 137–163, 2001.
[18] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
[19] J. Ham, Y. Chen, M. M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 492–501, 2005.
[20] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.
[21] R. Woolson, "Wilcoxon signed-rank test," *Wiley encyclopedia of clinical trials*, pp. 1–3, 2007.
[22] S. Lakhina, S. Joseph, and B. Verma, "Feature reduction using principal component analysis for effective anomaly–based intrusion detection on nsl-kdd," 2010.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on April 07,2021 at 14:40:08 UTC from IEEE Xplore. Restrictions apply.