# CLASSIFICATION BHP FLOODING ATTACK IN OBS NETWORK WITH DATA MINING TECHNIQUES

V.N. UZEL[1] and E. SARAÇ EŞSİZ[2]

[1]Adana Science and Technology University, Adana/Turkey, vnuzel@adanabtu.edu.tr
[2]Adana Science and Technology University, Adana/Turkey, esarac@ adanabtu.edu.tr

***Abstract* - Today, almost everything is done through networks. Especially, Networks are widely used for transportation of data. Various methods are used to move the data from one place to another. One of these methods is Optical Burst Switching (OBS). When carrying data in OBS, some of the threats may be encountered as a result of security shortcomings. Some of these threats are Spoofing, Replay Attack, Circulating Burst Header Attack and Burst Header Packet (BHP) Flooding Attack. Detection of threats is difficult but it is very important to our safety. Therefore, using Machine Learning (ML) methods to detect threats will give us flexibility, time and accuracy. In this study, we will classify BHP Flooding Attack data that have four class labels with ML methods. Our class labels are as follows: Misbehaving-Block (Block), Behaving-No Block (No Block), Misbehaving-No Block (NB-No Block), and Misbehaving-Wait (NB-Wait). Methods used in classification are Decision Tree (J48), Logistic, Multilayer Perceptron (MLP), Random Tree (RT), Reduce Error Pruning (REP) Tree and Naive Bayes (NB). Since there are 22 properties in the data set, the results of feature selection are also examined using the same classification methods. As a result, J48 and RT have been found to achieve the best results with 100% accuracy.**

***Keywords* – Machine Learning, Data Mining, Network Attacks, Optical Burst Switching (OBS) Network, Burst Header Packet (BHP) Flooding Attack**

## I. INTRODUCTION

TOGETHER with developing technology, transmission of data through networks has become the center of our lives.

Various methods are used to transmit the data. The information can be carried by conventional methods, such as cables, or it can be transported by the new optical method. Optical fibers can transport data further away and have higher bandwidth than electrical cables. Optical method uses light to carry information by providing a point-to-point connection. With Wavelength Division Multiplexing (WDM) technology, bandwidth is divided into number of non-overlapping wavelength channels. Optical methods that use WDM technology include Optical Circuit Switching (OCS), Optical Packet Switching (OPS), and Optical Burst Switching (OBS). OCS [1] is not suitable for intensive internet traffic. OPS [1] is flexible and has efficient bandwidth but there is a buffering problem. OBS [1] has huge bandwidth, lower error rates and security advantages. OBS combine the good aspects of OCS and OPS and cover their gaps. Comparisons of these technologies are given in Table 1.

Table 1: Comparison of Switching Technologies.

| FEATURES | OCS | OPS | OBS |
|---|---|---|---|
| Traffic Adaptability | Low | High | High |
| Bandwidth Utilization | Low | High | High |
| Buffering | No | Yes | Yes |
| Latency | High | Low | Low |
| Overhead | Low | High | Low |

In OBS, a burst is a data packet which can have variable length. Burst have two components: control and payload. The control packet carries the header information. The payload is the actual data transmitted. Firstly, Burst Header Packet (BHP) establish a path from source to destination. Then data is sent from this path. While data is being sent, it can be attacked due to lack of security. Some of these attacks are Spoofing, Replay Attack, Circulating Burst Header Attack and BHP Flooding Attack.

Spoofing [2] is used for accessing restricted files and information by the hackers. Hackers can access information easily by taking the IP address of a trusted network. The system assumes that it comes from a reliable source and accepts the packet exchange.

In Replay Attack [3], an attacker detects a data transmission and delayed or repeated this transmission fraudulently. For example, a user enters a website and logs in into his/her account with a password, then the website opens a session to the user. If an attacker intuits the session, attacker can login the user account with that session.

In Circulating Burst Header Attack [4], more than two compromised nodes organize for an attack. One of them acts as a master, and the others are slaves. For example, there are 'n' compromised nodes, one of them will be a master, (n-1) will be slaves. Slave nodes are listed in 1 to n-1. A node can only forward the BHP to the next node. A BHP can only transmitted to the destination through the master node. Therefore, a BHP should be transmitted from first node to the master node one by one. In this attack, network resources are wasted and prevented from being used by the new burst.

In BHP Flooding Attack [5], multiple copies of transmitted

BHP are created when any optical node is seized by the attackers. Along with the generated copies, a lot of BHPs are transmitted to the next node. So the next node tries to allocate space for fake BHPs. As a result, the resources can’t accept a valid BHP when it arrives.

In this study, BHP Flooding Attacks will be classified using Machine Learning (ML) methods for detection of attacks, to ensure network security.

The rest of the paper is organized as follows: in the second section a brief overview of the related work is presented. The dataset and the applied methods are described in the third section. The fourth section presents and discusses the experimental results. Finally, section five concludes our study.

## II. RELATED WORK

Rajab et al. [6] use the same dataset with our study. They use the Decision Tree as a classifier. In addition, feature selection is applied. They use the chi-square method as the feature selection and CFS method to verify the chi-square. First, classes are separated as misbehaving and behaving and they reached 93% accuracy rate. Then, the same dataset is divided into 4 subgroups as Misbehaving-Block (Block), Behaving-No Block (No Block), Misbehaving-No Block (NB-No Block), and Misbehaving-Wait (NB-Wait). With this method they reach 87% accuracy rate.

In a different study that are used the same dataset with ours ten-fold cross validation is applied while separating the dataset [7]. It means that the dataset is divided into 10 pieces and one of them is used as testing data, others are used as training data. Dataset is labeled with 4 classes such as our study. They used Naïve Bayes, Bayes Net, Decision Tree, and their suggested Rule Model classifiers. Naïve Bayes and Bayes Net classifiers reached 69% and 85% accuracy rates, respectively. Rule-model and Decision Tree classifiers reached over 98% accuracy rates.

Kavitha et al. [8] use the same dataset with our study. Ten-fold cross validation is applied. They used Decision Table, JRIP, OneR, PART-m, ZeroR, Naïve Bayes and Bayes Net classifiers. PART-m gave the best result with 100% accuracy in 0.02 seconds. Same dataset is classified with Decision Stump, Hoeffding Tree, J48, LMT and REP Tree. The best result of this classification is LMT with 100% but it has taken 4.59 seconds.

Villaluna et al. [9] use NSL-KDD and KDD99 datasets together for classifying attacks on the network. The datasets have 5 classes such as: Normal, DoS, Probe, U2R, and R2L. Fuzzy Logic, Artificial Neural Network and Fuzzy Neural Network are used as classifiers and the attack detection rate for the three algorithms are 94.84%, 98.51%, and 98.60% respectively. Also, accuracies of each algorithm are 89.74%, 96.09%, and 96.19% respectively.

Ormani et al. [10] used the NSL-KDD dataset, they divided the traffic into attack and normal. They proposed a fusion of ANN and SVM methods for classification. And they compared their results with ANN and SVM. As a result, their proposed method is achieved better classification results. The true positive rate results for ANN, SVM, and ANN + SVM are 79.56%, 79.27%, and 79.65%, respectively. When selecting flag and protocol features with feature selection, the result is reached 79.71%.

## III. MATERIALS AND METHODS

Firstly, the dataset is duplicated because of unbalancing data. Then it is classified with several classifiers. Also, feature selection is applied for better classification result. After that, results are compared with before feature selection and after feature selection.

### A. Dataset

In this study, “BHP flooding attack on OBS network dataset” is used from UCI dataset repository [11]. The dataset has 1075 instances and 22 features. The twenty-second feature is a class label. It has 4 class labels. These are Misbehaving-Block (Block), Behaving-No Block (No Block), Misbehaving-No Block (NB-No Block), and Misbehaving-Wait (NB-Wait). Firstly, the dataset is split into train and test datasets. Our dataset has an unbalanced distribution of class labels. To make it a balanced dataset, Block label is duplicated 4 times and No Block label is duplicated 3 times.

### B. Classifiers

J48 is a decision tree classifier that creates a binary tree with the help of information entropy. After the tree is built, it can be used to assign class labels to each tuple in the test dataset [12].

The Artificial Neural Network (ANN) [13] is a model inspired by the human brain and the nervous system. An ANN can have several layers. The first layer is called the input layer and the last layer is called the output layer. The middle layers are called hidden layers. Each layer contains a certain number of neurons connected by synapses. Multilayer Perceptron (MLP) [13] is a type of ANN and uses back propagation to train the network.

The Naïve Bayes [14] algorithm is a simple probabilistic classifier which uses the Bayes theorem. It computes probabilities by counting the frequencies of attribute values for each class in a given training dataset. The algorithm assumes all attributes are independent given the value of the class variable.

Logistic [15] is based on statistical results like other classifiers. It may be misleading because the name is a regression, but it does not predict continuous values. It is suitable for binary classification.

Random Tree (RT) [16] builds the largest tree and has the lowest performance among all type of trees. It considers a set of k- randomly chosen attributes to split on at each node. It performs no pruning.

Reduce Error Pruning (REP) Tree [16] is a fast decision tree algorithm. It does reduced-error pruning and considers all of attributes. As in the C4.5 Algorithm, this algorithm handles missing values by segmenting the corresponding samples.

### C. Feature Selection

Feature selection is a preprocessing step for machine learning methods. Aims of feature selection are reducing dimensionality, removing irrelevant data and increasing learning accuracy.

In Correlation-based Feature Selection (CFS) [17], a feature is important if it is highly relevant for classification or it is irrelevant with other features. Filter method is used for selecting features. CFS uses the correlation between features.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Environment

WEKA [18] data mining tool is used for feature selection and classification. Netbeans [19] is used for other preprocessing implementations like splitting the dataset into train and test.

### B. Evaluation Measures

Four measures are used for classification metrics. These are Accuracy, Precision, Recall, and F-measure [20]. Confusion matrix that is shown in Table 2 is used to compute these four measures.

Accuracy is the percentage of correctly classified samples in the test dataset. Precision is the ratio of true positives to all positively labeled samples. Recall is the ratio of true positives to all positive samples in the test dataset. F-measure is the harmonic mean of precision and recall. Equations 1, 2, 3, and 4 describes how to compute these four measures.

Table 2: Confusion Matrix

| | | **Predicted** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Actual** | **Positive** | True Positive(TP) | False Negative(FN) |
| | **Negative** | False Positive(FP) | True Negative(TN) |

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$\mathrm{Pr}ecision = TP/(TP + FP) \quad (2)$$

$$\mathrm{Re}call = TP/(TP + FN) \quad (3)$$

$$F_1 = 2(\mathrm{Pr}ecision + \mathrm{Re}call)/(\mathrm{Pr}ecision * \mathrm{Re}call) \quad (4)$$

### C. Experimental Evaluation and Results

J48, Logistic, MLP, NB, RT and REP Tree are used as classifiers. CFS is used for feature selection.

The aim of this study is detecting BHP flooding attack in OBS network. The dataset has 4 class labels and 3 of them are misbehaving, one of them is behaving. There are Block, No Block, NB-No Block, NB-Wait. Classification results without feature selection are given in Table 3.

Table 3: Classifier Results before CFS

| **Classifiers** | **Accuracy** | **Precision** | **Recall** | **F-measure** |
|---|---|---|---|---|
| **J48** | 100% | 1.000 | 1.000 | 1.000 |
| **Logistic** | 89.35% | 0.892 | 0.894 | 0.892 |
| **MLP** | 95.83% | 0.964 | 0.958 | 0.957 |
| **NB** | 81.48% | 0.822 | 0.815 | 0.810 |
| **RT** | 90.74% | 0.930 | 0.907 | 0.905 |
| **REP Tree** | 97.22% | 0.972 | 0.972 | 0.972 |

According to Table 3, without feature selection, J48 has the best accuracy rate. Classification results with CFS feature selection are given in Table 4.

Table 4: Classifier Results after CFS

| **Classifiers** | **Accuracy** | **Precision** | **Recall** | **F-measure** |
|---|---|---|---|---|
| **J48** | 100% | 1.000 | 1.000 | 1.000 |
| **Logistic** | 86.57% | 0.862 | 0.866 | 0.859 |
| **MLP** | 89.35% | 0.892 | 0.894 | 0.892 |
| **NB** | 82.41% | 0.796 | 0.824 | 0.783 |
| **RT** | 100% | 1.000 | 1.000 | 1.000 |
| **REP Tree** | 93.98% | 0.941 | 0.940 | 0.940 |

With CFS feature selection method, 10-Run-AVG-Bandwith-Use and Flood Status features are selected from 22 features. According to Table 4, with CFS, J48 and RT are the best accuracy rates, performance of NB and RT are increased, J48 is remained the same and others are decreased.

## V. CONCLUSIONS AND FUTURE WORK

Various problems may be encountered while moving the data through networks. One of them is a BHP flooding attack. The aim of this study to detect BHP flooding attacks with Machine Learning method and analyze effects of feature selection on classifiers.

It is observed that, Machine Learning methods can be used for detection of BHP flooding attacks with high accuracy rates. And feature selection has no significant effects on classifier performance for this dataset.

## REFERENCES

[1] P.K. Chandra, A.K. Turuk and B. Sahoo, "Survey on Optical Burst Switching in WDM Networks", *Fourth Intemational Conference on Industrial and Information Systems*, Sri Lanka, December 2009.

[2] K. Jindal, S. Dalal and K.K. Sharma, "Analyzing Spoofing Attacks in Wireless Networks," *2014 Fourth International Conference on Advanced Computing & Communication Technologies*, Rohtak, India, Feb. 2014.

[3] A. Jesudoss and N.P Subramaniam, "A Survey On Authentication Attacks and Countermeasures In A Distributed Environment," *Indian Journal of Computer Science and Engineering (IJCSE)*, Vol. 5, Apr-May 2014.

[4] N. Sreenath, K. Muthuraj, and G. Vinoth, "Threats and Vulnerabilities on TCP/OBS Networks," *2012 International Conference on Computer Communication and Informatics (ICCCI -2012)*, Coimbatore, INDIA, Jan 2012.

[5] K. Muthuraj and N. Sreenath, "Secure Optical Internet: An Attack on OBS node in a TCP over OBS network," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, Vol. 1, November-December 2012.

[6] A. Rajab, C. Huang and M. Al-Shargabi, "Decision Tree Rule Learning Approach to Counter Burst Header Packet Flooding Attack in Optical Burst Switching Network," *Optical Switching and Networking*, Vol. 29, pp. 15-26, July 2018.

[7] R. Alshboul, "Flood Attacks Control in Optical Burst Networks by Inducing Rules using Data Mining," *IJCSNS International Journal of Computer Science and Network Security*, Vol. 18, February 2018.

[8] S. Kavitha1, M. Hanumanthappa and A. Syrien, "Evaluation of Optical Burst Switching (OBS) Using Various Classification Techniques," *National Conference On Contemporary Research and Innovations in Computer Science (NCCRICS)*, Dec. 2017.

[9] J.A. Villaluna and F.R.G. Cruz, "Information Security Technology for Computer Networks through Classification of Cyber-Attacks using Soft Computing Algorithms," *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, Dec. 2017.

[10] T. Omrani, A.Dallali, B.C. Rhaimi, J. Fattahi, "Fusion of ANN and SVM Classifiers for Network Attack Detection," *18th international conference on Sciences and Techniques of Automatic control & computer engineering*, Monastir, Tunisia, December 21-23, 2017.

[11] archive.ics.uci.edu, 'Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network Data Set', 2017. [Online]. Available:https://archive.ics.uci.edu/ml/datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+Network. [Accessed: 16 - July - 2018].

[12] I. Jenhani, N.B. Amor and Z. Elouedi, "Decision trees as possibilistic classifiers," *International Journal of Approximate Reasoning*, Vol. 48, pp. 784-807, 2008.

[13] M.W. Gardnera, S.R. Dorlinga, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmospheric Environment*, Vol. 32, pp. 2627 – 2636, 1998.

[14] D.D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval", *Machine Learning: ECML-98*, Vol. 1398, pp. 4-15, 1998.

[15] C.J. Peng, K.L. Lee and G.M. "Ingersoll, An Introduction to Logistic Regression Analysis and Reporting," *The Journal of Educational Research*, Vol. 96, pp. 3-14, September – October 2002.

[16] D.L.A. AL-Nabi1 and S.S. Ahmed, "Survey on Classification Algorithms for Data Mining: (Comparison and Evaluation)," *Computer Engineering and Intelligent Systems*, Vol. 4, 2013.

[17] S. Vanaja and K.R. Kumar, "Analysis of Feature Selection Algorithms on Classification: A Survey," *International Journal of Computer Applications*, Vol. 96, June 2014.

[18] cs.waikato.ac.nz, [Online]. Available: https://www.cs.waikato.ac.nz/ml/weka/. [Accessed: 16 – July 2018].

[19] netbeans.org, [Online]. Available: https://netbeans.org/. [Accessed: 16 – July 2018].

[20] D.M.W. Powers, "Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies*, Vol. 2, pp. 37-63, 2011.