**CHITTAGONGUNIVERSITY OF ENGINEERING AND TECHNOLOGY**

**OFFICE OF THE MEMBER SECRETARY OF THE COMMITTEE FOR HIGHER STUDIES AND RESEARCH (CHSR) CUET, CHITTAGONG.**

**(THESIS PROPOSAL)**

**Application for the approval of M.Sc.Engg. Thesis Proposal**

Status: Part Time

1. **Name of the Student** : Muhammad Kamrul Hossain Patwary

 **Roll No**: 15MCSE043P **Session: 2015-16**

2. **Present Address** : Flat C, Building no 7, Building type B, Port east
 colony, Port, Chittagong, Bangladesh.

3. **Name of the Supervisor** :Dr. Md. Mokammel Haque

4. **Name of the Department** :Computer Science and Engineering

5. **Date of First Enrolment** :23/06/2016

6. **Tentative Title** :Malicious Attack Detection and Blocking

 Using Semi-supervised Machine Learning Technique

 in Optical Burst Switching Network

7. **Background and Present state of the Problem:**

An Optical Network (ON) is commonly used to transmit data from a source to its destination, using light via an optical fibre medium. In contrast to traditional networks, ON features efficient quality performance indicators such as bandwidth and speed. Thus, it is a preferable option for Internet infrastructure. In order to make use of the huge bandwidth of ON, Optical Burst Switching (OBS) was proposed in [1] as being the next generation of optical switching technology. Once it has obtained the User Datagram Protocol (UDP) packets, an OBS network will assemble the packets from the clients at the edge nodes (ingress node) into a data burst (DB) and a burst header packet (BHP), which will be transmitted in advance to preserve the network resources required before the DB is actually sent. Security and QoS performance issues are two important issues that need to be addressed to guarantee OBS network reliability.



Figure 1: An optical burst switching network data flow

The work in [2,3,4] have discussed malicious attacks in Layer-1 in optical networks, and proposed solutions to reduce the impact of Layer-1 attacks. However, Layer-3 flooding attack still pose open threats, and attackers can exploit this, making an ingress node (source node) overload (flood) the network with BHPs that reserve the resources without transmitting the actual DB [5]. Besides, fake/malicious DB may cause denial of service in the network. It is important, therefore, to ensure that prevention of BHP flooding attacks and fake/malicious DB attacks, is a high priority in OBS. These can potentially severely reduce the performance of the entire network, eventually causing a Denial of Service (DoS).Despite the many advantages of the OBS network, such as resiliency, bandwidth efficiency, as well as its economic benefits, QoS and security can become issues, with consequences including burst loss as a result ofBHP flood attacks [6, 7]. Attacks of this type are reliant on the flooding approach, which has been examined in traditional DoS against the TCP protocol [8].



Land Attack:

The compromised node copies the BHP and transmits back to the source and to the intended destination. Due to the fact that the attack is on a split capable node, the data also gets split and reaches both intended and unintended nodes thereby wasting resources.

Burstification attack:

A particular node could be compromised thus changing the value of the size of the assembled bursts at the BHP. The increased burst size value could push the egress node to check the value of the same during disassembly. Due to the attack, if the value is not comparable, the burst could be mistaken for another burst forcing the receiving node to ask for retransmission of the burst.



Figure 2: BHP flooding attack model

A limited number of studies exist in relation to dealing with and preventing issues caused by BHP flood attackswithin OBS networks (such as, in [9, 7, 6, 10]). For example, [9] proposed a flow filtering architectureoperating at the optical layer to filter out BHP flood attacks at an early stage. The filtering process is performedthrough a comparison between the offset time included in the BHP and the actual delay between it and theassociated DB. In [7], the authors examined the issue of DoS within the resources’ reservation protocols. Thiscountermeasure adopts the method of using optical code words, filtering out false BHPs, and identifying thecompromised source node within the network. In [6], meanwhile, the authors proposed a prevention method,built by gathering statistical data from packets used to detect a BHP flood attack in TCP within an OBSnetwork. In [10], a new security model was proposed and developed to be integrated within an OBS core switcharchitecture to prevent BHP flooding attacks. Using a countermeasure security model enables the core switch ofthe OBS to classify the ingress nodes according to their behaviour, in addition to calculating the amount ofreserved resources not being used. A malicious node, which causes a BHP flooding attack, can be blocked usingthis model until the risk has disappeared.

The issues with these methods, however, is that they heavily rely ondomain experts’ rules to label misbehaving edge nodes rather than the methods used as seen in machine learning(ML). In ML, detecting misbehaving edge nodes can be implemented based on hidden correlations discoveredfrom historical data, and are related to the sending edge nodes’ behaviours such as delay time, packet drop rate,received bytes among others; thus, automatically classifying edge nodes to the appropriate class in thepreliminary risk stage in order to limit the risks of BHP flooding attacks. Conventional methods are limited inmany cases, especially when the computer network encounters change in the behaviour of sending nodes’ datatransmission, therefore needing further verification of their analysis results. ML application can be useful especially in cross-layer settings, where data analysis at physical layer, can trigger changes at network layer e.g. monitoring Bit Error Rate (BER) can trigger changes at network layer in routing, spectrum and modulation format assignments. Overall, ML techniques offerpromising solutions, potentially effective in dealing with BHP flood attacks in computer network security,offering automated classification systems derived with minimal human involvement. This indeed relaxes theextra care, domain experience and time limitations associated with most existing solutions.This study will examine the problems of DoS resulting from BHP flooding attacks and fake/malicious DB attacks. In BHP flooding attack, legitimate BHPsare prevented from preserving network resources for legitimate DBs, whereas in fake/malicious DB attack, bogus DBs are sent to flood the network. The ML architecture we have developedfeatures sets of beneficial learning gathered from past simulations, carried out using a reduced number offeatures including the bandwidth used, the average packet drop rate, as well as the average delay time persecond.One of the most encouraging data analysis methods used by researchers for prediction is ML, featuringintelligent techniques in order to complete a specific task, usually linked to knowledge building or isolatingpatterns which are concealed. Little research exists around the classification of edge nodes in preventingBHP flooding attacks within OBS networks.

Rajab et. al. [10] proposed a new security model to prevent BHP flooding attacks. The model allows the OBS core switch to classify the ingress nodes based on their behavior and the amount of reserved resources that are not being utilized. But they did not use any machine learning, also they did not discuss about fake/malicious burst send by malicious nodes. Rajab et al. [11] proposed a supervised, decision tree-based architecture that extracts rules. The results showed that the rules derived accurately classify 93% of the BHP flooding attacks into either Behaving (B) or Misbehaving (M) classes. Moreover, the rules can further classify the Misbehaving edge nodes(M) into four sub-class labels with 87% accuracy. But they did not discuss about empty burst send by malicious nodes. Musumeci et al. [12] suggested that, after learning from a batch of available past samples, other types of algorithms, in the field of semi-supervised and/or unsupervised ML, could be implemented to gradually take in novel input data as they are made available by the network control plane.Mata et al. [13] suggested that, artificial intelligence will continue playing an important role in supporting emerging transmission technologies like space division multiplexing, multimode/multicore fibers and advanced modulation formats and constellation shaping. So, research should be increased on these fields. Coulibaly et al. [14] proposed a solution to address Data Burst Redirection (DBR) Attack in OBS networks. The solution is designed based on Rivest-Shamir-Adleman (RSA) public-key encryption algorithm. But they only detected fraud BHP and did not classify the compromised nodes based on their behavior.Chen et al. [15] outlined several security concerns in OBS networks like orphan bursts, malicious burst headers etc. They introduced key based authentication of burst headers and confidentiality of data bursts

This study will propose a new series of rules using the decision tree method to prevent the risks of the BHP flood attack and fake/malicious DB attack problem. Firstly, the selected ML method will build a binary classification model, determining the edge nodes into two classes (Behaving node and Misbehaving node). Hundreds of simulation runs will be used to collect the data to build the binary models, gathering the various attributes associated with how the edge nodes perform. Next, the classification models will be improved by splitting the Misbehaving class into further sub-class labels, establishing a priority procedure for data transmission from the edge nodes. The proposed classification rule models will be evaluated using different metrics, measuring the overall performance of this approach. Lastly, we shall compare the result of our work with other similar works in this field.

8. **Objectives with specific aims and possible outcomes:**

The major objectives of this thesis are:

1. To find better solution to BHP flooding attack.
2. To find a solution to Data Burst flooding attack.
3. To find a Semi-supervised machine learning architecture to prevent flooding attack.

9. **Outline of Methodology:**

We propose a security model to prevent BHP flooding attacks and fake/malicious DB attacks, using machine learning model.Our aim is to classify the nodes based on their behavior and the amount of reserved resources that are notbeing utilized.A malicious node that causes such attack will be blocked by the developed model until the risk disappears.

We will design a semi-supervised architecture. The reason for choosing semi-supervised architecture is that, the vast majority of existing studiesadopting ML at the networking level use offline supervisedlearning methods, i.e. assume that the ML algorithms aretrained with historical data before being used to take decisionson the field. This assumption is often unrealistic for opticalcommunication networks, where scenarios dynamically evolvewith time due, e.g., to traffic variations or to changes in thebehavior of optical components caused by aging. Moreover, inpractical assets, it is difficult to collect extensive datasets duringfaulty operational conditions, since networks are typicallydimensioned and managed via conservative design approacheswhich make the probability of faults negligible (at the priceof under-utilization of network resources). We thus envisagethat, after learning from a batch of available past samples,other types of algorithms, in the field of semi-supervisedML, could be implemented to graduallytake in novel input data as they are made available by thenetwork control plane. Moreover, scarce attention has so farbeen devoted to the fact that different applications might havevery different timescales over which monitored data show observable and useful pattern changes (e.g., aging would make component behaviour vary slowly over time, while trafficvaries quickly, and at different time scales (e.g., burst, daily,weekly, yearly level).



Figure 3: Proposed semi-supervised architecture of machine learning model

Algorithm to consider:

For supervised learning:

* Random forest
* Neural network
* Kernel SVM
* Gradient Boosting tree
* K-NN
* Naïve bayes
* Logistic regression

For Unsupervised learning:

* KMeans

Machine Learning Tools:

Scikit-learn: It is a free machine learning library for the Python programming language.

Dataset:

* Dataset from University of California Irvine (UCI) repository, namely, “Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network Data Set.”
* Dataset from NCTUns network simulator

Evaluation:

* The security model will be implemented, tested and verified using NCTUns network simulator
* We shall see if it is effective in countering BHP flooding attacks as well as in providing the network resources to the legitimate nodes.
* We shall compare our work with some other similar work.

10. **References:**

1. C. Qiao, M. Yoo, Optical burst switching (OBS) - a new paradigm for an optical Internet, Journal of High Speed Networks. vol. 8 no.1 (1999) pp. 69-84
2. N. Skorin-Kapov, J. Chen, and L. Wosinska, “A new approach to optical networks security: Attack-aware routing and wavelengthassignment,” IEEE/ACM Trans. Netw., vol. 18, pp. 750–760, Jun. 2010
3. J. Zhu, B. Zhao, W. Lu, and Z. Zhu, “Attack-aware service provisioning to enhance physical-layer security in multi-domain EONs,”J. Lightw. Technol., vol. 34, pp. 2645–2655, Jun. 2016.
4. J. Zhu, B. Zhao, and Z. Zhu, “Leveraging game theory to achieve efficient attack-aware service provisioning in EONs,” J. Lightw.Technol., vol. 35, pp. 1785–1796, May 2017.
5. J. Turner, Terabit burst switching, Journal of High Speed Networks. vol. 8 (1999) pp. 3-16.
6. A.W. Moore, D. Zuev, Internet traffic classification using Bayesian analysis techniques, In: ACM SIGMETRICS PerformanceEvaluation Review. vol. 33, no. 1, 2005, pp. 50-60.
7. M. Sliti, N. Boudriga, BHP flooding vulnerability and countermeasure, Photonic Network Communications. 29(2), (2015) pp.198-213.
8. W.M. Eddy, TCP SYN Flooding Attacks and Common Mitigations. RFC 4987 (2007).
9. M. Sliti, M. Hamdi, N. Boudriga, A novel optical firewall architecture for burst switched networks, in: Proc. 12th Intl. Conference onTransparent Optical Networks (ICTON). (2010) pp. 1-5.
10. A. Rajab, C.T. Huang, M. Alshargabi, and J. Cobb, Countering Burst Header Packet Flooding Attack in Optical Burst SwitchingNetwork, In: International Conference on Information Security Practice and Experience. Springer International Publishing. Nov 16(2016) pp. 315–329
11. A. Rajab, C.T. Huang, M. Alshargabi, Decision tree rule learning approach to counter burst header packet flooding attack in Optical Burst Switching network, In Optical Switching and Networking 29 (2018): 15-26
12. F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, M. Tornatore. A Survey on Application of Machine Learning Techniques in Optical Networks. arXiv preprint arXiv:1803.07976. 2018 Mar 21.
13. J. Mata, I. de Miguel, R. J. Durán, N. Merayo, S.K. Singh, A. Jukan, M. Chamania, Artificial intelligence (AI) methods in optical networks: A comprehensive survey. In Optical Switching and Networking, (2018).
14. Y. Coulibaly, A. A. I. Al-Kilany, M. S. A. Latiff, G. Rouskas, S. Mandala,M. A. Razzaque, Secure burst control packet scheme for Optical Burst Switching networks. In Broadband and Photonics Conference (IBP), 2015 IEEE International (pp. 86-91).
15. Y. Chen, P. K. Verma, Secure optical burst switching: Framework and research directions, In IEEE Communications magazine,(2008) 46(8).