GPON ATTACKS AND ERRORS CLASSIFICATION

Adrián Tomašov
Doctoral Degree Programme (1st), FEEC BUT
E-mail: xtomas32@vutbr.cz

Supervised by: Tomáš Horváth
E-mail: horvath@feec.vutbr.cz

Abstract: This paper focuses on various types of attacks and errors in an activation process of Gigabit-capable passive optical networks. The process sends messages via Physical Layer Operation Administration and Maintenance header field inside the transmitted frame. An exemplar network communication is captured by a special hardware-accelerated network interface card capable of processing optical signals from passive optical networks. The captured data is filtered of irrelevant parts and messages and correctly formatted into a suitable shape for a neural network. The filtered data is divided into small sequences called time windows and analyzed using a recurrent neural network-based on Gated recurrent unit cells. A new neural network model is designed to classify sequences into several categories: additional message, missing message, error inside (noisy) message, and message order error. All of these categories represent a certain type of attack or error. The proposed model can distinguish message sequences into these categories with high accuracy resulting in revealing a possible attacker or drift from protocol recommendation.

Keywords: Activation Process, GPON, GRU, Recurrent Neural Network, PLOAM

1 INTRODUCTION

Passive optical networks are currently the most promising solution providing broadband internet connectivity, video streaming, etc. The distributional network consists only of passive components (e.g., fiber cables, splitters) and has a tree-like topology, where Optical Line Termination (OLT) is the root and leaves are devices called Optical Network Unit (ONU). The messages are broadcasted into the network and received by all members of communication. Thus, it is necessary to keep confidentiality of transmitted data and synchronization due to multiple access.

Gigabit-capable Passive Optical Networks (GPON) protocol is defined by a set of recommendations by Inter-national Telecommunication Union (ITU) [1]. A recommendation is not as strong as standard, so device vendors can modify this protocol to match their requirements without proper description. This allows attackers to send malicious messages of various formats through the specific parts of the frame without deciding whether the message is a part of the protocol or a malicious message sent by an attacker. With a specific targeted attack, the attacker would prevent clients from communicating through the network, change the priority of data flow and proceed with different denial of service or abuse them for his benefit.

One of the weak points is the activation process. This process is defined through a finite state machine, and a transition between states is triggered by specific messages located in a Physical Layer Operation Administration and Maintenance (PLOAM) field in a frame header. This process's main purpose is to synchronize newly activated ONUs and prepare them for transmission. It also handles error states (loss of signal, loss of framing), resulting in restarting this process and temporary outage.

This paper focuses on the development of a classification model based on Machine Learning (ML) algorithms. The model can learn and recognize messages and communication patterns from exemplar

communication, which represents trusty communication. The model analyzes message sequences in GPON and classifies every malicious or strange behavior different from the exemplar one. This type of analysis is specific and requires a particular type of neural network capable of time series evaluation.

Neural networks are not a novelty in ML; however, they were limited by computational resources until recently. Today technology supports hardware acceleration allowing the development of complex machine learning models applied in various fields.

Time series analysis is a specific ML problem requiring a modified model for a certain data format. Models based on recurrent neural networks are often used for such analysis. These models consist of recurrent cells capable of sharing internal state/states to themselves through sequence processing. Xueqi Zhang et al. (2020) [2] evaluate and compare traditional predictive methods against a recurrent neural network based on Long Short-Term Memory (LSTM) cells. These models predict an amount of ambient town noise. The best accuracy has the model with LSTM layers.

Xin Wei et al. (2021) [3] compare neural networks based on LSTM, Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (SimpleRNN) and based on several dense layers. These models predict pore-water pressure depending on the amount of rainfall water in a certain time segment. LSTM and GRU layer-based models have clearly the best accuracy in such prediction. At the same time, authors refer to GRU cell effectively, which requires about 40% fewer resources compared to LSTM based model.

2 ANALYZED DATA

The most important part of each ML project used for data analysis is the learning dataset. Even a perfect neural network design would have poor performance, accuracy, and generality with insufficient dataset quality. Therefore, it is necessary to keep enough attention to the data itself, which are used as examples for the ML model. The first step is to analyze the dataset using mathematical and statistical tools to reveal any possible hidden error or bias. This section describes data capturing and data pre-processing algorithms and the final data shape suitable for the neural network. In the end, there are function descriptions used for generating attacks and errors applied to captured communication.

2.1 DATA CAPTURE

This work is based on data captured from a real GPON network. The frames are captured by a custom Network Interface Card (NIC) containing Field Programmable Gate Array (FPGA) capable of converting signals from the optical domain into a suitable format for machine processing. The converted data from the optical network are sent into a data server, which parses the frames. The extracted information is being correctly formatted and stored in a database. This work is based on the database data, which are exported in JavaScript Object Notation (JSON) format.

The analysis is focused on the ONU activation process. Therefore, during the data capturing, ONUs are randomly disconnected from a power supply, or attenuation of optical connection is randomly changed. These events randomly execute the whole activation process several times by each ONU, which provides a wide range of the activation process usage and suitable data for the neural network.

2.2 DATA PRE-PROCESSING AND FORMAT

The captured dataset is filtered of irrelevant messages and message fields before preparing as input for the neural network. Considering the work is focused on attacks and errors in GPON activation process, the most relevant data are located in PLOAM field in the frame header. Other parts and message headers are deleted because they do not have any influence on the activation process. Consequently,

the rest of the data[1] is being inspected further. Messages with ID 11 are 99.98% of all captured messages and deleted from the dataset. Otherwise, they would cause enormous distortion, and trained models would not learn any useful patterns.

The output of the filtration process is a sequence of 13 bytes messages. This sequence is too long to be used directly as an input of the neural network model. Therefore, the sequence is divided into smaller segments called time windows. The number of messages in segments is 30, which is long enough to contain the whole activation process for several ONU units and short enough to be used as input for the learning process. After division, there is a list of time windows with shape 13×30 bytes. In the end, the dataset is normalized into float numbers by dividing by 255.

2.3 ERROR/ATTACK GENERATION

Several functions are used to generate data into the learning dataset containing searched errors. Each function's input is a cloned instance of captured messages, which is being modified by each function directly. The modified sequences are generated time windows with a specific width suitable for the neural network input. The generated error datasets are:

- **Deleted message error** dataset is generated by deleting important messages of the activation process. Important messages have these IDs: 4, 8, 10, and 18. This function iterates through each important message and deletes all occurrences of that message in a separate instance of captured message sequences. The output of this function is almost four times longer sequence compared to the input sequence.
- **Additional message error** generation process is also focused on messages of the activation process. This function iterates through message: 4, 8, 10, and 18, and for each occurrence of the current message, this function doubles the message. The output sequence is more than four times longer than the input sequence.
- **Order error** is generated using a function, which creates time windows according to the given format from the input sequence. The next step inverts messages in time windows, which clearly generates corrupted message sequences.
- **Invalid message error** containing invalid messages is generated by adding Gaussian noise into the captured communication with a mean equal to 0 and standard deviation equal to 0.2. The generated message sequence is different from the captured communication and contains various errors simulating a malicious device or possible attack.

3 NEURAL NETWORK MODEL

Inspiration for the model is natural language processing (text), speech recognition (voice), and various other time series classification models. Most of them use neural networks based on convolutional or recurrent layers. Based on similar researches and experiments during this study, the proposed neural network model is based on GRU recurrent layer, which has enough tools for long and complex sequences [4]. The GRU cell design is simpler compared to LSTM cell, so it requires fewer hardware resources, but still can solve the exploding/vanishing gradient problem gently.

3.1 DESIGN

Considering the data characteristic and format, there are two GRU recurrent layers with 64 cells at the beginning of the neural network followed by two dense layers. GRU layers can analyze time

[1] The rest of the data contains PLOAM messages only.

series, which is exactly the problem this model investigates. The remaining dense layers conclude the outcomes of recurrent layers and classify the sample into the correct category. The number of neurons in the last dense layer is equal to the number of classified categories.

3.2 Learning

An exemplary dataset learns the model with generated sequences described in Section 2.3 and uses supervised learning. There is no early stop function used during learning, but the model's current state is saved after each learning epoch. The batch size parameter is set to 128. This process is executed exactly 200 epochs to reveal the tendency of loss and accuracy metrics when the network is being over-fitted. The optimizer used for learning is *Adam*. The loss function used for an error evaluation is `sparse categorical cross-entropy`, which can convert expected labels from numbers to `one-hot encoding` and evaluate the network error.

The learning dataset is obviously imbalanced, which is visible in data methods described in Section 2.3. Thus, class weights are computed to eliminate an error caused by the imbalanced dataset. These weights are used only during learning, and they scale the loss value for each neuron according to category differently. The weights are calculated using `compute_class_weight` from `sklearn.utils.class_weight` module. The whole dataset is divided into three disjunctive sets in a ratio of 70:15:15 to fairly evaluate the quality and other parameters of the final model.

3.3 Results Evaluation

The proposed neural network can recognize attacks with high accuracy. Considering the time window size, it can recognize possible attacks or errors through the PLOAM field in the ongoing communications. The network quality is evaluated using confusion matrices showing to which category each sample is classified. The confusion matrix evaluated on the testing dataset is visible in Figure 1. The figure proves that the recurrent neural network's proposed architecture accurately classifies message sequences into all categories. However, there are some false predictions of `additional` and `missing` dataset samples, possibly because these datasets may contain sequences of massages without modified parts. Thus, these sequences are the same as sequences in the exemplary datasets and should be deleted from datasets representing attacks and errors.

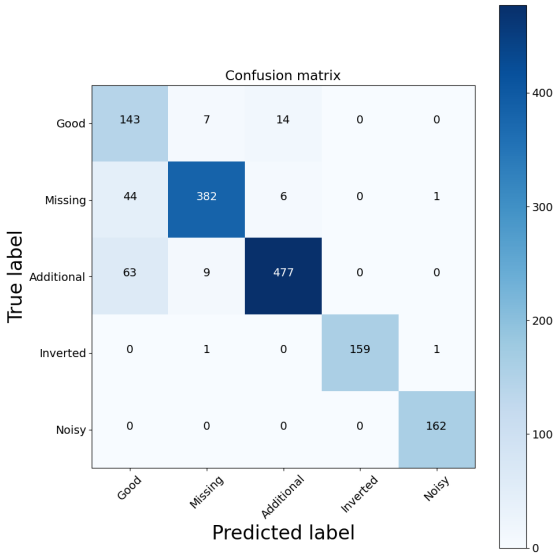

Figure 1: The confusion matrix of test set evaluation.

4 CONCLUSION

This work investigates the neural network design used for the analysis of PLOAM messages. Based on similar works focusing on time series analysis, the proposed model is based on layers with GRU cells. These cells are more effective than LSTM, but with almost the same accuracy of classification.

The exemplar communication is captured in the real GPON network in the laboratory environment. This data contains many useless parts, fields, and message types, and the data, if filtered to consists of related information only. The remaining data generate other datasets containing various types of errors or attacks on the activation process. All datasets are formatted into the time windows suitable for neural network learning.

The recurrent neural network's proposed design shows very high accuracy in analyzing the activation process, which is visible in the confusion matrix of the test dataset evaluation and the accuracy evolution of the validation dataset during learning.

This work is one of the first in the field of machine learning analysis of GPON protocol and reveals possibilities of developing an Intrusion Prevention System (IPS) and Intrusion Detection System (IDS) based on machine learning models. Furthermore, most importantly, this reveals the possibility of real-time inspection of the communication inside a passive optical network. This gives many options for enhancing this work in several topics:

- Anomaly detection models based on unsupervised learning, especially auto-encoders.
- Evaluation possibility of a convolutional network for messages sequence analysis.
- Enhance data capture process to gain ONT Management Control Interface messages.

ACKNOWLEDGEMENT

The research described in this paper was financed by a grant from the Ministry of the Interior of the Czech Republic, Program of Security Research, VI20192022135, PID VI3VS/746 for "Deep hardware detection of network traffic of next generation passive optical network in critical infrastructures".

REFERENCES

[1] ITU-T. Gigabit-capable passive optical networks (g-pon): Transmission convergence layer specification. [online], 01 2014.

[2] Xueqi Zhang, Meng Zhao, and Rencai Dong. Time-series prediction of environmental noise for urban iot based on long short-term memory recurrent neural network. *Applied Sciences*, 10(3):1144, 2020.

[3] Xin Wei, Lulu Zhang, Hao-Qing Yang, Limin Zhang, and Yang-Ping Yao. Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. *Geoscience Frontiers*, 12(1):453 – 467, 2021.

[4] R. Dey and F. M. Salem. Gate-variants of gated recurrent unit (gru) neural networks. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1597–1600, 2017.