Hitachi Energy
HITACHI
Inspire the Next

Study Committee D1

Materials and Emerging Test Techniques

Paper D1-PS1-10397

MACHINE LEARNING ALGORITHM TRAINED BY THE DUVAL PENTAGONS: A SIMPLIFIED DUVAL PENTAGONS APPROACH

LUIZ CHEIM, PHD

HITACHI ENERGY

Motivation

- Difficulty to classify multiple and historical DGA data from online sensors using complex Duval Pentagon Geometry
- Train a Machine Learning Algorithm using conventional Duval Pentagon and create a simple function to classify any combination of 5 Hydrocarbons used in the Pentagon method

Method/Approach

- The DGA classification in the Pentagons is done by the location of a geometrical centroid originated by the composition of 5 combustible gases, as illustrated in Figure 1 below.
- Create thousands of artificially generated centroids to train the Machine Learning Algorithm and obtain a single function that eliminates the need for the complex geometry of the Pentagons.

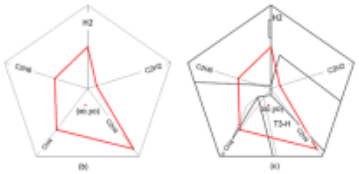

Figure 1 – Pentagon conventional DGA classification

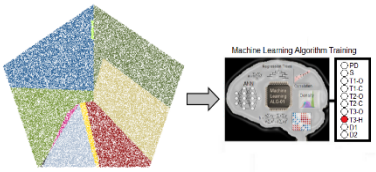

Figure 2 – Training Machine Learning Algorithm based on artificially generated centroids

Objects of investigation

- Artificially generated DGA samples based on a multitude of centroids randomly generated inside the Pentagons.
- Real DGA samples from a large number of transformers

Experimental setup & test results

- Multiple Machine Learning Algorithms were trained, and their performance compared until the best model was selected (best accuracy).
- A large dataset of real DGA samples from operating transformers tested to demonstrate Machine Learning DGA classification accuracy.

Discussion

- Artificially generated data are acceptable to train the Machine Learning algorithm for the fact that the whole Pentagons region can be covered by such approach and any real case DGA sample will necessarily fall in a location where the artificial data is located.
- Although many combinations of artificially generated data may not correspond to real life cases of DGA, any real-life case of DGA will necessarily correspond to at least one artificially generated data.

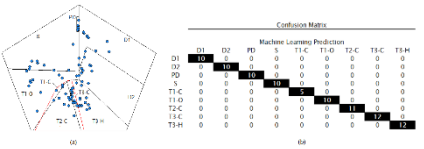

Figure 3– Machine Learning Accuracy (100 test cases)

Conclusion

- The work presented in this paper demonstrated that a Machine Learning algorithm was successfully trained by the Combined Duval Pentagons and was capable of correctly classifying many DGA samples (never seen by the algorithm during the training phase) with an accuracy above 99%.
- Now, if DGA measurement errors and the issue of misclassification due to boundary conditions in the multiple regions of the Pentagons are considered, such accuracy seems to be adequate with a substantial statistical confidence. Finally, the trained algorithm becomes a simple function, easy to implement in any platform, avoiding the need to implement complex logic and geometries for classification of DGA samples in the Combined Pentagons.