**Paper Summarized: Characterization of Encrypted and VPN Traffic using Time-relate Features by Draper-Gil et. al.**

**Set Objectives of the paper:** To study the effectiveness of flow-based time-related features to detect VPN traffic and to characterize encrypted traffic into different categories, according to the type of traffic e.g., browsing, streaming, etc They also claimed to reduce overhead by reducing the set of features. However, procedures followed to accomplish features reduction was not explicitly reported.

**Their Contributions**

1. Flow and Features Generation: They used SCXFlowMete, a Java application written by this researcher for the generation of dataset that we shall be using for out experiments.
2. Use of Weka to implement C 4.5. decision tree and KNN algorithms to classify generated datasets using 10 folds cross validation.

**Their Results and Discussion**

1. They found out that there is a direct relationship between flow-timeout value and performance of the classifiers. With 15 seconds as the best threshold. **This finding contradicts 600s timeout that literature have been using. (NB: we need to find about the correctness of this claim and reason for such contradiction)**

**Our Project relevant Takeaways:**

**Our suggested improvements over what was done**

1. To compare if there is a statistically significant difference between the results of more than two algorithms in the classification of VPN traffic using multiple flow-based time related features.
2. Application of feature selection and analysis of its effect on the algorithms’ performances.

**A Question to answer:** Are their other features apart from time related?

**Some Literature Building ideas**

Traffic classification can be categorized based on its final purpose: associating traffic with encryption (e.g., encrypted traffic), protocol encapsulation e.g., tunneled through VPN or HTTPS); according to specific applications, (e.g., Skype), or according to the application type (e.g., Streaming, Chat),

Generally speaking, the classification of network traffic falls mainly into two categories: flow-based classification, using properties such as flow bytes per second, duration per flow, etc. and packet-based classification, using properties such as size, inter-packet duration of the first (or *n*) packets, etc.