Practical and Configurable Network Traffic Classification using Probabilistic Machine Learning

Jiahui Chen
University of Utah

UUCS-20-007

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

30 April 2020

Abstract

Network traffic classification that is generally applicable and highly accurate is extremely valuable for many network security and management tasks. A flexible and easily configurable classification framework is ideal so it can be customized for use in many different networks. In this thesis we propose a highly configurable and flexible machine learning traffic classification method that relies only on statistics of sequences of packets to distinguish known or approved traffic from unknown traffic. Our method is based on likelihood estimation, provides a measure of certainty for classification decisions, and can classify traffic at adjustable certainty levels. Our classification method can also be applied in different classification scenarios, each prioritizing a different classification goal. We demonstrate how our classification scheme and all its configurations perform well on real-world traffic from a high performance computing network environment.