# Unsupervised Two-Class & Multi-class Support Vector Machines for Abnormal Traffic Characterization

[Extended Abstract]

Angelos K. Marnerides[†], Dimitrios P. Pezaros[†], Hyun-chul Kim[‡], and David Hutchison[†]

[†]Computing Department, Infolab21, Lancaster University, Lancaster, UK, LA1 4WA

{a.marnerides, dp, dh}@comp.lancs.ac.uk

[‡]School of Computer Science and Engineering, Seoul National University, Seoul, South Korea, 151-742

hyunchulk@gmail.com

## ABSTRACT

Although measurement-based real-time traffic classification has received considerable research attention, the timing constraints imposed by the high accuracy requirements and the learning phase of the algorithms employed still remain a challenge. In this paper we propose a measurement-based classification framework that exploits unsupervised learning to accurately categorise network anomalies to specific classes. We introduce the combinatorial use of two-class and multi-class unsupervised Support Vector Machines (SVM)s to first distinguish normal from anomalous traffic and to further classify the latter category to individual groups depending on the nature of the anomaly.

## Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General –*Security and protection*

## General Terms

Algorithms, Security, Measurement

## Keywords

Anomaly classification, machine-learning

## 1. INTRODUCTION

The categorisation of observed traffic flows has proved a great challenge due to the high processing requirements imposed by the high-volume of aggregate network data and the high-speed of backbone links. The real-time classification of attack-flows in particular, has also never been fully met [4]. The Support Vector Machine (SVM) is a promising linear machine learning-based classification scheme that sets strong foundations towards a real-time automated classification framework. The use of supervised SVM has showed encouraging results, achieving a 98% of classification accuracy on very high volumes of backbone traffic traces [3]. At the same time, researchers have successfully applied the unsupervised two-class version of SVM to distinguish normal from abnormal traffic [5]. However, neither effort has focused on the exact classification of particular type of network anomalies (e.g. DDoS vs. flashcrowd, etc.).

In this paper, we exploit research conducted on SVMs, Kernel methods (KMs), and in particular quadratic and semidefinite programming that has formulated a version of SVM employing two-class and multi-class unsupervised classification [1], to perform real-time traffic analysis. We propose to apply the unsupervised multi-class SVM technique in the domain of network anomaly diagnosis and further detailed attack-specific traffic classification.

## 2. REAL-TIME CLASSIFICATION

Our proposed framework consists of a real-time measurement-based classifier that initially separates normal from abnormal traffic and further classifies in detail the anomalous traffic flows within the abnormal space. In particular, the classifier operates on the distributions of selected flow features using unsupervised learning, alleviating the need to possess pre-labelled data in order to classify new events. Therefore, by mitigating the need for offline statistical analysis, such framework puts strong candidacy for real-time traffic classification.

The overall system is composed by two unsupervised SVM classifiers. As it is graphically illustrated in Figure 1, we firstly use the categorisation capabilities provided by the two-class unsupervised SVM and subsequently we employ the multi-class unsupervised SVM. The rationale behind this two-phase strategy is to initially filter out the classified normal traffic in order for the multi-class classifier to concentrate on a subset of traffic and classify specific attack flows. Initial flow classification is based on seven packet header characteristics; the IP source/destination addresses, the transport protocol, the transport source/destination ports, the mean packet inter-arrival time and the size of the first ten packets, as also suggested in [4][3]. During the second stage, only four packet header features are used; the source/destination IP addresses and transport source/destination ports. This latter set of packet features has been shown to exhibit distinct deviations in their distributions where those denote particular attacks[4].

### 2.1 Separation of Normal & Abnormal Traffic

The first phase of flow classification aims at separating normal from abnormal traffic behaviour, and it is achieved using the two-class unsupervised SVM.

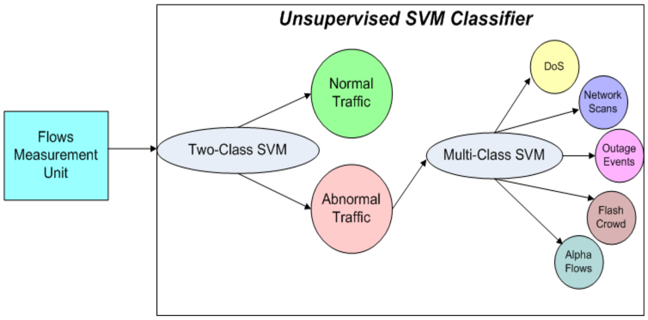

**Figure 1: A high-level representation of the classification framework**

The unsupervised (mostly two-class) SVM aims to find a labelling that results in a large margin classifier (i.e. the maximum margin of separation between two classes) in an *n*-dimensional feature space [1]. This is in contrast to the supervised SVM [6] that finds the margin classifier based on the labels of data [1]; rather it is achieved by a two-class clustering approach. The rationale in this case is to initially find a label that if it was about to run, a SVM would obtain a maximal margin over all possible labels [2]. Despite its high computational requirements, such approximation may be computed by a semidefinite program which would enable the re-expression of the margin issue in terms of the label kernel matrix, instead of directly referring to the cluster labels [1]. Therefore, in our case of having the distributions of flow features as the unlabelled data $x^1,...,x^n$, we desire to solve for a binary labelling $y \in \{-1,+1\}^n$ which will potentially lead to a maximum margin. The justification behind the choice of the particular values assumed by *y* is outside the scope of this paper, and can be found in [2].The semidefinite approach will subsequently produce and use the label kernel matrix $M = yy^T$. The advantage provided by the re-expression of the optimal (i.e. maximal) margin is that the inverse square margin $\gamma^{*-2}$ is a convex function of $M$ and this allows to directly express the optimal margin within $M \in \{-1,+1\}^{nxn}$. Therefore, in our case the input features in the **-1** class are denoted as normal traffic and the rest as abnormal, similar to the classification in [5]. Figure 2 shows the separation of the two classes in a two-dimensional plane, emphasising the margin boundaries denoting the actual SVMs.

## 2.2 Abnormal Traffic Classification

Following the separation of normal and abnormal traffic, we employ the multi-class unsupervised SVM to further classify attack flows. We use this latter algorithm to re-express the supervised multi-class SVM [1] using semidefinite semantics [1]. The semidefinite expression constructs new equivalence indicator matrices that allow the allocation of multiple labels to multiple margin classifiers [1]. By using the four-feature tuple mentioned in section 2, we target the extraction of meaningful clusters each one denoting a particular class of abnormality (e.g. DDoS vs. network scan). The input for this second stage will be the group of the distributions of the four selected flow features that were classified within the **+1** class during the first stage.

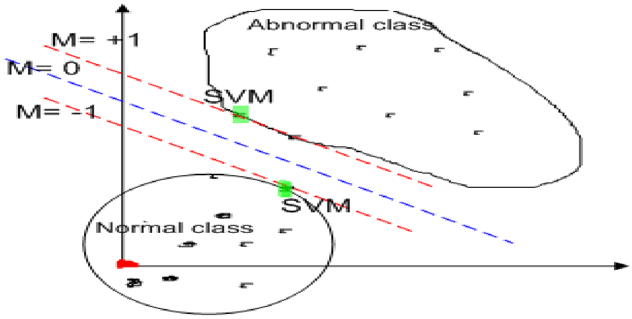

**Figure 2: Geometrical 2D interpretation of two-class unsupervised SVM**

# 3. CONCLUSIONS & FUTURE WORK

The primary goal of the proposed measurement-based framework is to enable accurate real-time traffic classification. Its design employs a foundation for specifically categorising hard to record anomalies under an unsupervised persona. The results by its supervised counterpart [3] strengthen the argument that the unsupervised mode will achieve high accuracy rates on a real-time scheme. Our objective is to initially test the proposed framework using offline analysis of large data sets and to subsequently deploy a real-time network-wide classification system.

# 4. ACKNOWLEDGEMENTS

This work was supported with the kind support of the Korea Research Council of Fundamental Science & Technology, the IT R&D program for Fundamental Technologies for the Future Internet (MKE/IITA 2007-F-038-02) and the EU FP6 Autonomic Network Architecture (ANA) project (FP6-IST-27489).

# REFERENCES

[1] Crammer K., Singer Y., On the Algorithmic Interpretation of Multiclass Kernel-Based Vector Machines, Journal of Machine Learning research (JMLR) 2, Vol. 2, pp. 265-292, 2001

[2] Joachims T., Transductive, Inference for Text Classification Using Support Vector Machines, in International Conference on Machine Learning (ICML'99), Bled, Slovenia, June 27-30, 1999

[3] Kim H., Claffy K. C., Fomenkov M., Barman D., Faloutsos M., Lee K., Internet Traffic Classification Demystified: Myths, Caveats and the Best Practises, ACM CoNEXT'08, Madrid, Spain, December 9-12, 2008

[4] Lahkina, A., Crovella, M., Diot, C., 2005, Mining Anomalies Using Traffic Feature Distributions, ACM SIGCOMM 2005, Philadelphia, PA, August 22-26, 2005.

[5] Li K., Teng G., Unsupervised SVM Based on p-kernels for Anomaly Detection, IEEE Int. Conference on Innovative Computing, Information and Control (ICICIC), Beijing, China, August 30 – September 1, 2006

[6] Vapnik V. N., The Nature of Statistical Learning Theory, Springer, ISBN 0-387-98780-0, New York, 1999

[7] Xu L., Schuurmans D., Unsupervised and Semi-supervised Multi-class Support Vector Machines, National Conference on Artificial Intelligence (AAAI-05), Pittsburgh, PA, July 9-13, 2005,