# Network Anomaly Classification by Support Vector Classifiers Ensemble and Non-Linear Projection Techniques

**Emiro de la Hoz[1,3] - Andrés Ortiz[2]**

**Julio Ortega[1] - Eduardo de la Hoz[1,3]**




(1) Computer Architecture and Technology Department. CITIC. University of Granada. 18060 Granada, Spain.
(2) Department of Communications Engineering. University of Málaga. 29071. Málaga, Spain.
(3) Systems Engineering Program. Universidad de la Costa. Barranquilla, Colombia.

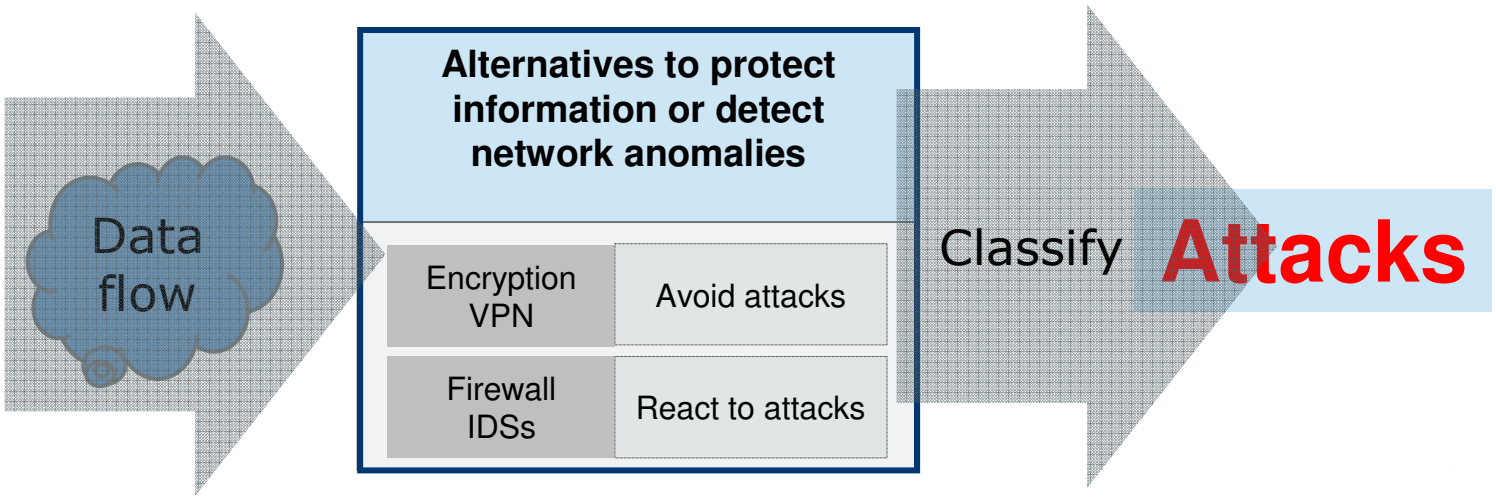
# Overview

- Categories of attacks
- Situation
- Classification system proposed
- Data preprocessing stage
- Feature selection stage
- Dimensionality reduction using PCA, Kernel PCA and Isomap
- Classification using Support Vector Classifiers (SVC) ensemble
- Experimental setup
- Experimental results
- Conclusions
- Future works

# Categories of attacks

1. Denial of Service Attack (DoS)
2. Probing Attack (PROBE)
3. User to Root Attack (U2R)
4. Remote to Local Attack (R2L)

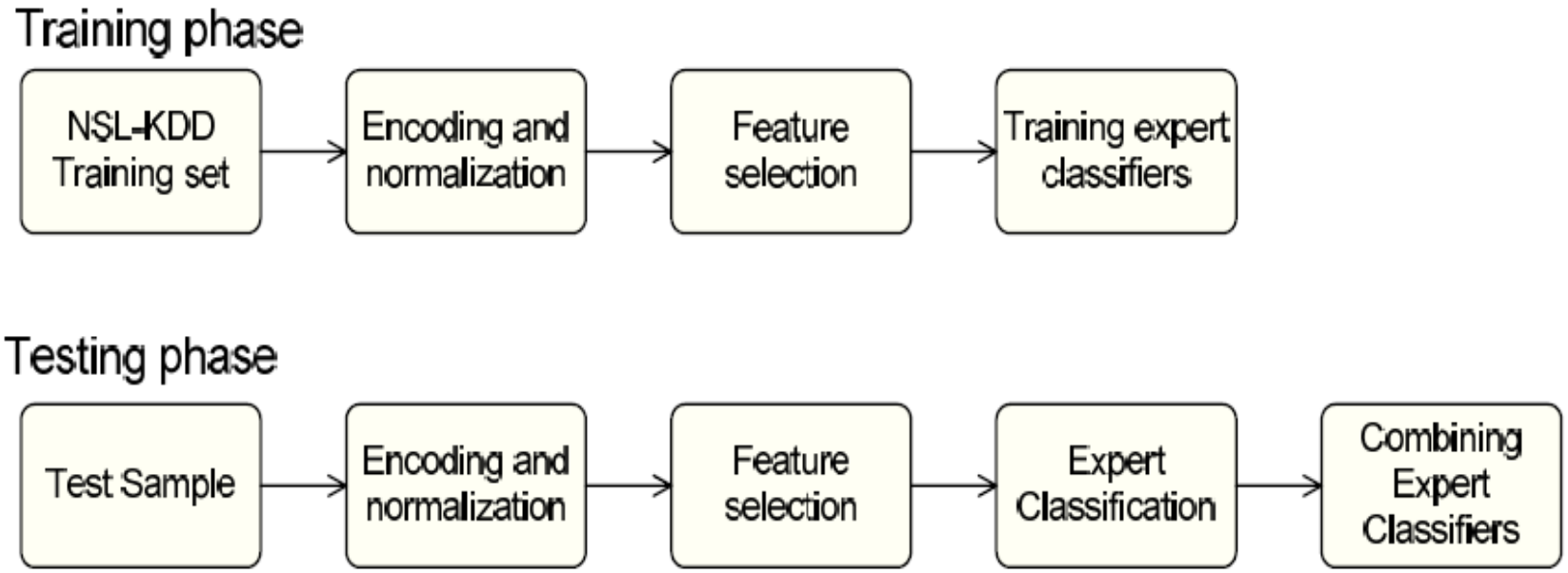
# Situation




**Improve the classification rate of network attacks and categorize such attacks using IDS**

# Classification system proposed

**Anomaly detection method using a support vector classifiers and non linear projection techniques.**



# Data preprocessing stage (encoding and normalization)

- We have used the NSL-KDD dataset, where each record consists of 41 features which can be symbolic and binary.

- We have mapped each connection type into a subspace Rd.

- Data normalization ensures that all the features are in the same scale. In this work, **cont**inuous variables are normalize**d** to “zero **m**ean” and “unity variance”.

# Feature selection stage

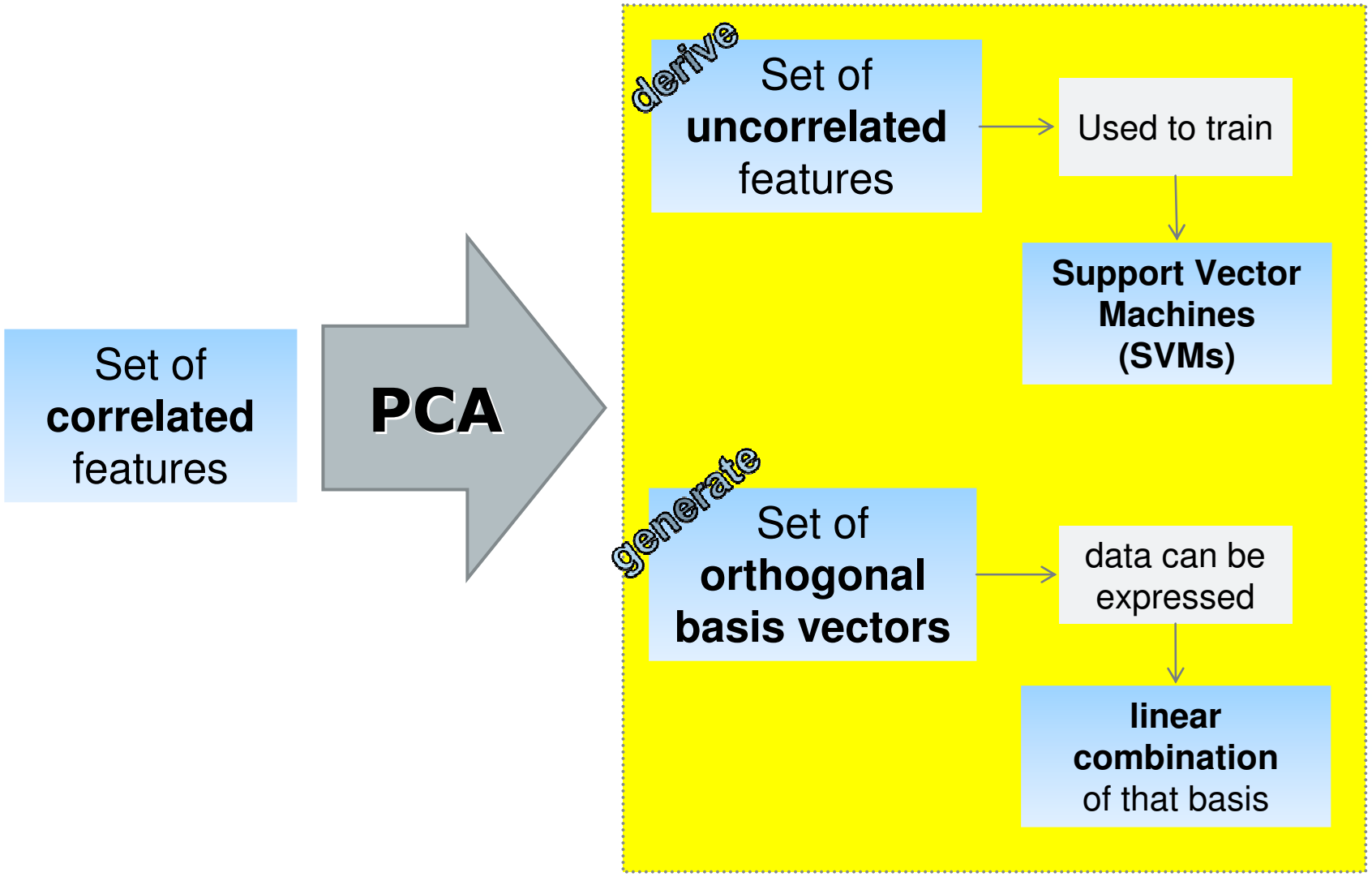
- We used the Fisher Discriminant Ratio (FDR), to preselect the features with the highest discriminatory power for each connection type.

$$FDR = \frac{(\mu_i - \mu_j)^2}{\sigma_i + \sigma_j}$$

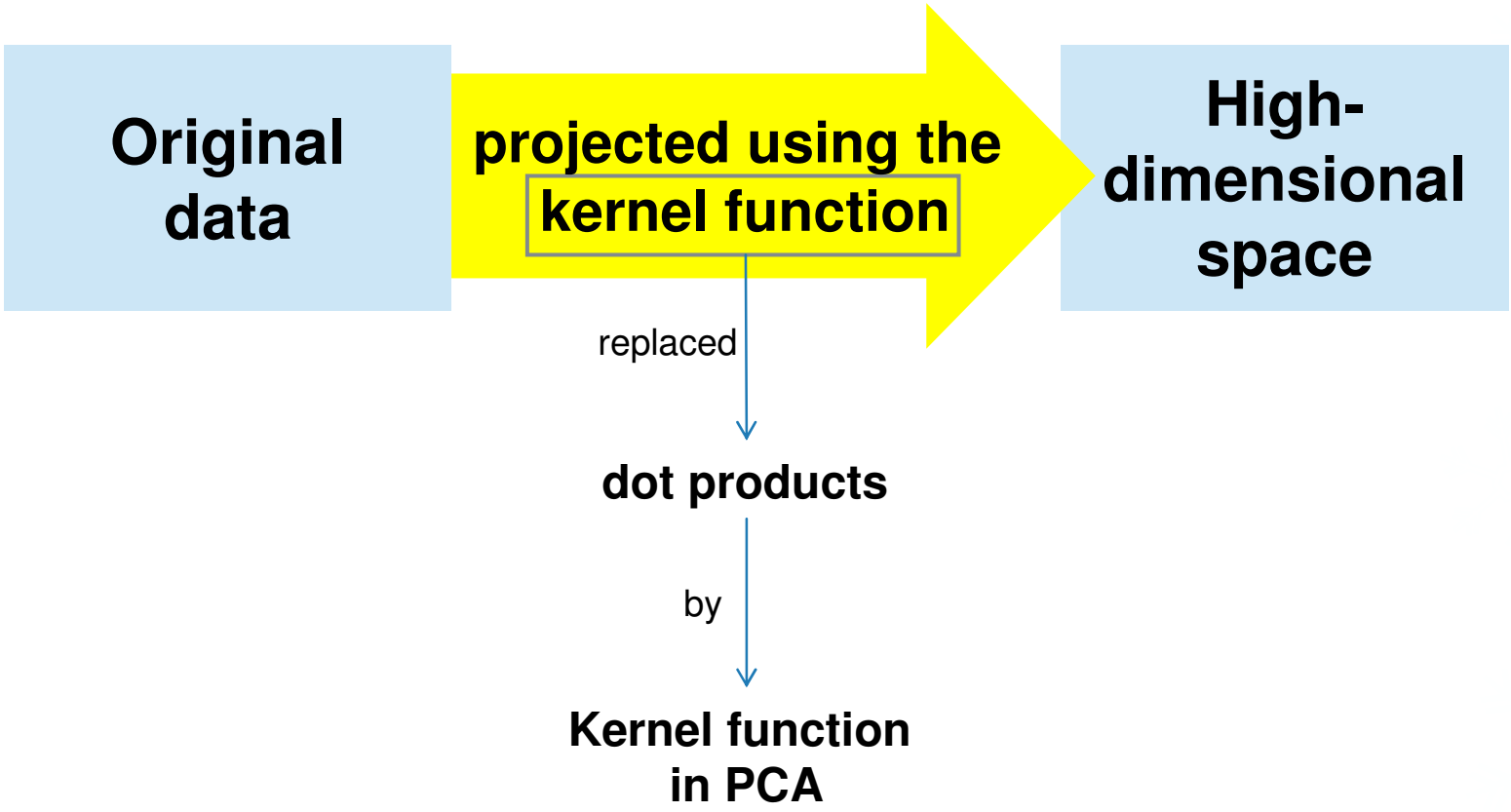
Equation FDR for the 2-class separation case

- We used techniques both linear dimensionality reduction (PCA) and nonlinear (Kernel PCA and Isomap) to demonstrate the effectiveness of data-based embedded KDD99.

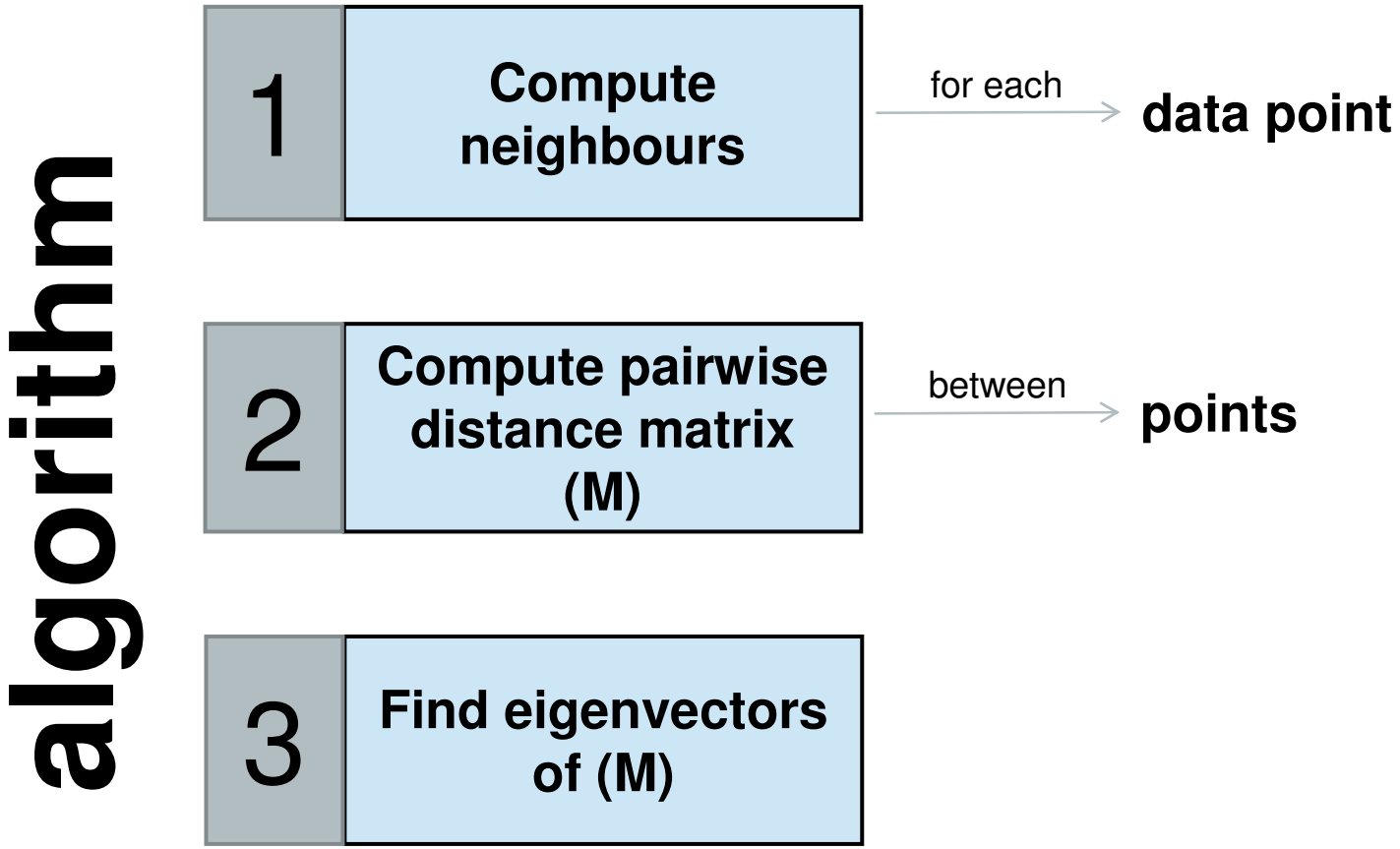
# Dimensionality reduction using PCA



# Dimensionality reduction using Kernel PCA



# Dimensionality reduction using Isomap



# Classification using Support Vector Classifiers (SVC) ensemble

Given a dataset {x1, ...xn} and its corresponding class labels {y1, ..., yn}, SVC's aim to solve the following optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

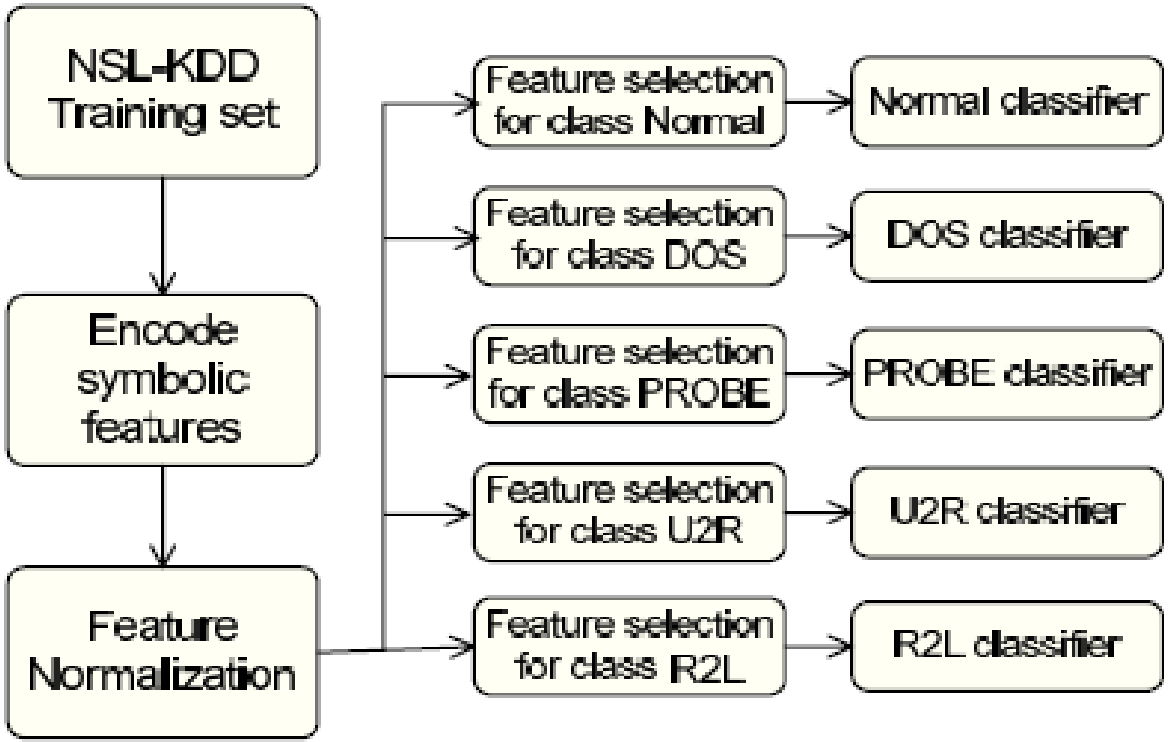
subject to constraint

$$y_i\left(w^T \phi(x_i) + b\right) \geq 1 - \xi_i,\ \xi_i \geq 0$$

φ = defines the kernel function transforming training vectors $\mathbf{X_i}$ into a higher dimensional space.

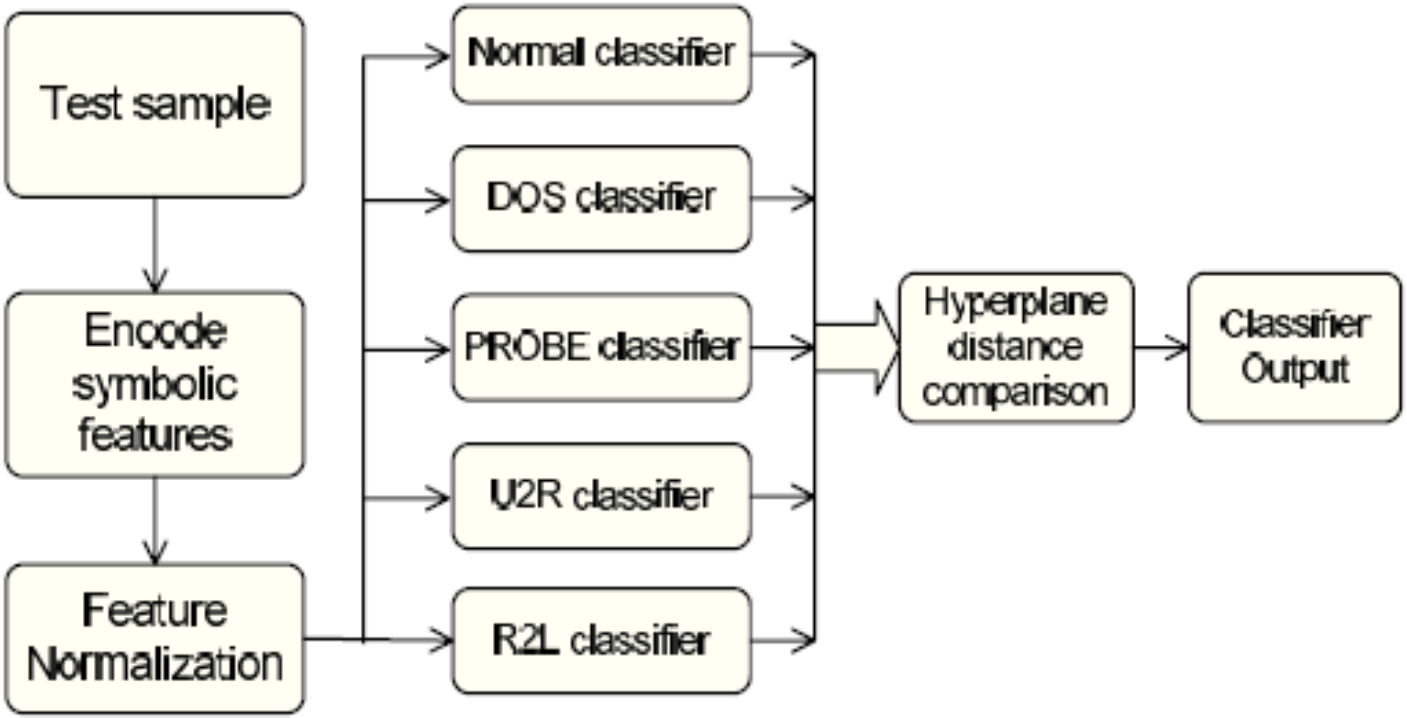
RBF kernel = defined as $K(x_i,x_j)=\exp(-\Upsilon\ ||x_i-x_j||^2)$, $\Upsilon>0$.

# Classification using SVC ensemble – Training Phase




Specific features computed from each attack type are used to train a binary classifier specialized in differentiate a specific attack from the rest.

# Classification using SVC ensemble – Testing Phase

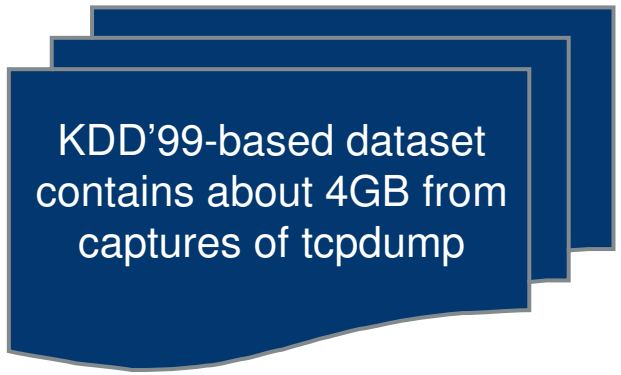



In the case of two or more experts classified the same sample as belonging to different classes, the class corresponding to the maximum distance to the SVC hyperplane is selected.

# Experimental Setup

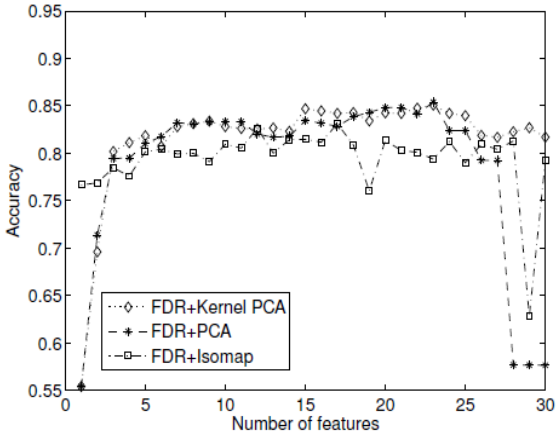
| | |
|---|---|
| Proposed method to detect network anomalies: | **cross-validation** |
| Training and testing data: | **NSL-KDD** |

**Dataset NSL-KDD**

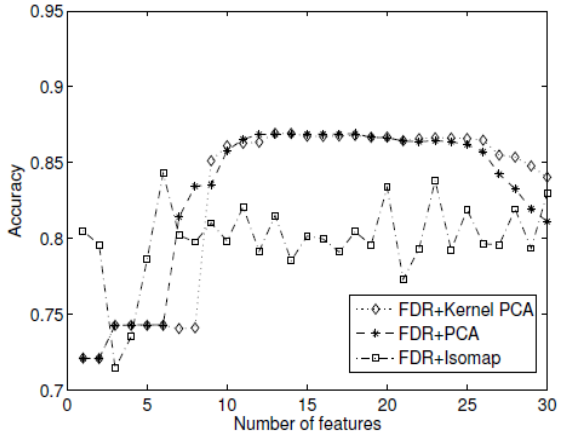



each connection contains 41 features

# Experimental Results

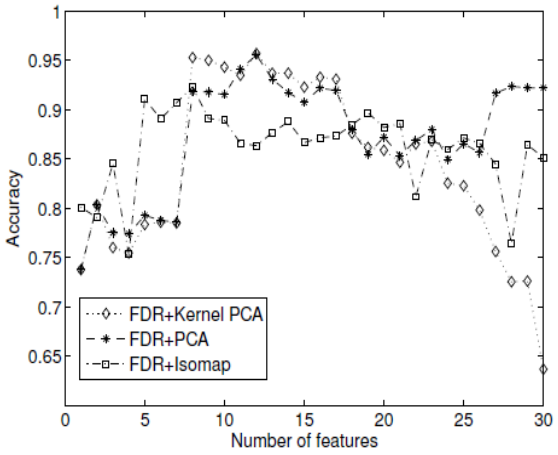



Normal connections

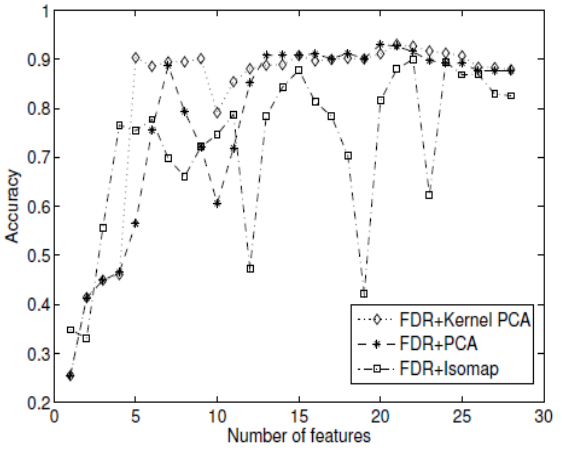



DOS

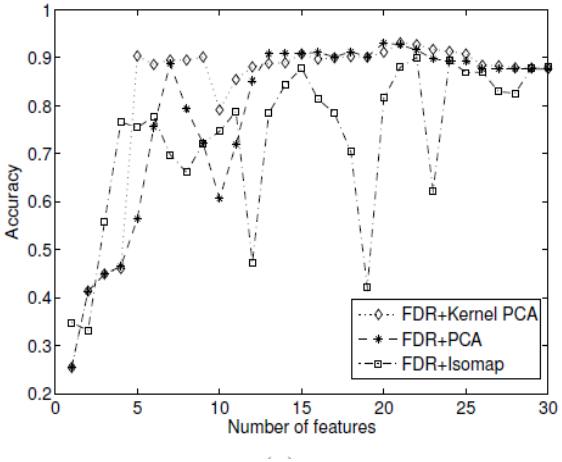
**Comparison among different linear and non-linear feature reduction techniques**




**Probe**



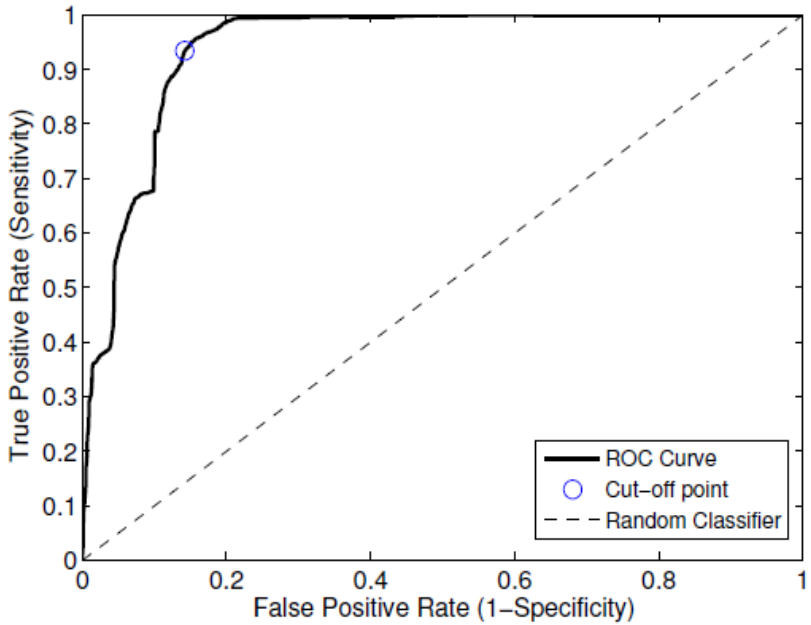

U2R




R2L

**The maximum accuracy is achieved using between 5 and 10 features.**

# Experimental Results




**ROC curve for normal/attack classification using KernelPCA as dimensionality reduction method**

| Method | Number of features | True Positive Rate | False Positive Rate |
|---|---|---|---|
| DM-Naïve Bayes [16] | 41 | 96.5 | 3.0% |
| Proposed method | 23 | 93.4% | 14% |
| Random Forest [18] | 41 | 80.67 | ** |
| Decision Trees [18] | 41 | 81.05 | ** |

** Data not provided by the author

**Attacks different types Tested separate**

Normal connections: 0.94
Attack DoS: 0.86
Attack Probe: 0.93
Attack U2R: 0.81
Attack R2L: 0.91

# Conclusions

- Presented a classification approach for **network anomaly classification** that combines nonlinear dimensionality reduction techniques and an **SVC ensemble** to build expert classifiers.
- Three different **dimensionality reduction techniques** have been used to assess their ability to generate discriminative features for each attack type.
- The maximum accuracy is achieved using between 5 and 10 features. Non-linear techniques (Kernel PCA and Isomap) perform **slightly better than** linear PCA.

# Future works

- Improve the method optimizing parameters in both dimensionality reduction and SVC classifiers.

- Improve the system of classification samples including more classifiers in the ensemble and optimized features for the two-class case.


Thanks!