# An autonomous labeling approach to SVM algorithms for network traffic anomaly detection

Carlos A. Catania[1], Facundo Bromberg[2], Carlos Garcia Garino[1]

[1] ITIC, Universidad Nacional de Cuyo, Mendoza, Argentina
{ccatania,cgarcia}@itu.uncu.edu.ar
[2] Dept. Sistemas de Información, FRM, UTN, Mendoza, Argentina
fbromberg@frm.utn.edu.ar

**Abstract.** In the past years, several support vector machines anomaly detection approaches have been proposed in the network intrusion detection field. The main advantage of these approaches is that they can characterize normal traffic when trained using a data set containing not only normal traffic but also possible attacks. Unfortunately, these algorithms seem to be accurate only when the normal traffic vastly outnumbers the numbers of attacks or anomalies present in the dataset.
This work presents an approach for autonomous labeling of normal traffic as a way of dealing with situations where class distributions do not present the required unbalance. The autonomous labeling process is made by SNORT, a misuse-based intrusion detection system. Experiments conducted on the 1998 DARPA dataset show the proposed autonomous labeling approach not only outperforms existing SVM alternatives but also obtains significant improvement over SNORT itself.

## 1 Introduction

In the past years network security has become a serious problem. In the early years of the Internet, the set of network protocols that support it worked reasonable well. However as the Internet grew, underlying security faults in those protocols were observed. Security faults in protocols such as ARP, TCP, TELNET, SMTP and FTP have caused most of known attacks against network data confidentiality, authenticity and availability. Currently all of these problems have been fixed, however new ways to develop attacks are discovered everyday.

Network managers must be well prepared in order to prevent network attacks, i.e., be informed about new vulnerabilities. For several years, intrusion detection systems (**IDS**) provided an invaluable help to network managers, becoming an integral part of any network security package.

In the intrusion detection field two different approaches can be observed: misuse detection and anomaly detection [1]. The idea behind misuse detection is to represent attacks in a form of a pattern or a signature in such a way that

even variations of these attacks can be detected. Based on these signatures, this approach detects attacks through a large set of rules describing every known attack. The main disadvantage of the signature based approach is its difficulty for detecting unknown attacks. The main idea of the anomaly detection approach is to build a statistical model for describing normal traffic. Then, any deviation from this model can be considered an anomaly, and recognized as an attack. Notice that when this approach is used, it is theoretically possible to detect unknown attacks, although in some cases, this approach can lead to a high false positive rate. This capacity to detect unknown attacks has been the cause of the increasing interest in developing new techniques to build models based on normal traffic behavior in the past years.

The anomaly detection approach has been a very active research topic inside the machine learning community and it has been the subject of many articles over the past years. One of the most successful approaches is based on the idea of collecting data only from network normal operation. Then, based on this data describing normality, any deviation would be considered an anomaly . Different techniques were proposed for characterizing the concept of normality [2]. In practice, however, it is difficult to obtain clean data to implement these approaches. Verifying that no attacks are present in the training data can be an extremely tedious task, and for large samples this is simply infeasible. On the other hand, if the data containing attacks is assumed clean, intrusions similar to the ones present in the training data will be accepted as normal patterns, resulting in an increment in the number of misdetections.

Recently, different authors proposed the use of support vector machines (**SVM**) for novelty detection [3,4,5] as an alternative approach for intrusion detection. One of the major advantages of this approach is that it is suitable for handling a training data set with not only normal traffic but also anomalies (i.e., attacks). Unfortunately, as was noticed by Eskin in [3], this works under the assumption that the number of normal traffic instances vastly outnumbers the number of anomalies, with a proportion of at least 98.5% of the training set being normal traffic.

This last assumption is not necessarily true in every situation. It is possible to find periods of time where the number of attacks present in traffic could easily outnumber normal traffic instances. This situation can also be observed in commonly used datasets for intrusion detection evaluation such as the 1998 DARPA dataset [6]. This dataset was provided by DARPA to the machine learning community in the context of the 1999 KDD Cup for evaluating different IDS approaches. Since its publication it has been widely used by many IDS researchers over the years. Interestingly, the 1998 DARPA class distribution does not exhibit the required unbalance. Moreover, it has many days or even weeks where the percentage of attacks raises up to 50%.

To deal with this and other unbalanced class distribution situations a novel approach is proposed. The idea is using SNORT [7], a very well known misuse signature-based IDS system, as an autonomous tool for labeling normal traffic.

The main hypothesis is that using SNORT may possibly reduce the presence of attacks in the traffic instances used for training, and consequently improving the performance of SVM for anomaly detection.

The rest of the work is organized as follows: in section 2 main characteristics of SVM for anomaly detection are briefly discussed, together with its application to the traffic network detection field. Then, in section 3, a new approach for autonomous labeling normal traffic is presented. In section 4 a set of experiments are conducted on the 1998 DARPA dataset for evaluating performance of the discussed approaches. Finally, conclusions and future work are provided in section 5.

## 2 SVM for anomaly detection

Since their introduction in the mid-1990s, *support vector machines* [8,9] have been widely used, being the subject of many articles on classification problems. SVM for anomaly detection is an extension of the core SVM ideas for classification problems. Traditional SVM approaches for classification uses as input training data consisting of a mixture of data labeled by both classes. In the intrusion detection problem this would consist of data labeled both as *attack* and *non-attack*. The model constructed by these approaches discriminates the input space in two infinite regions, one per class, usually using a hyperplane. In contrast, the main idea in SVM for anomaly detection [10,11] is to use as input a description of only the *normal* class of objects (*non-attack* in IDS), assuming the rest as *anomalies* (*attacks*). The model constructed by this approach discriminates the input space in a finite region containing the normal objects, while all the rest of the (infinite) space is assumed to contain the anomalies.

The SVM for anomaly detection variants appear in the literature of intrusion detection with different names, which could led to some confusion. In some cases they are referenced as SVM one-class algorithms. SVM for non supervised learning is another widely used name by some authors. Although, all of these names describe important characteristics of this kind of algorithms, in this work the term SVM for anomaly detection will be preferred.

Two major approaches were proposed to extend SVM for anomaly detection problems. One approach, proposed by Tax and Duin [10], is based on the idea of finding an hypersphere with center $c$ and minimal radius $R$ containing most of the *normal* data, discriminating all other data not in the sphere as *anomalies*. As in standard SVM approaches, the discriminating surface (the sphere), as well as the data, may be mapped into a feature space by some kernel function. Another approach proposed by Scholkopf [11] tries to separate the normal data points from the anomalies by finding the hyperplane that is maximally distant from the origin. When a RBF kernel is used, it was shown the two approaches converge to the same solution [12].

Due to space restrictions only a brief description of Tax's approach is provided. For a description of the hyperplane formulation the reader is referred to [11].

### 2.1 SVM based on the hypersphere formulation

The sphere formulation has an intuitive geometric idea: the normal data can be concisely described by a sphere enclosing the data in some high-dimensional feature space. A graphical example of this can be observed in Fig. 1. The presence of noise, i.e., incorrectly labeled training data, can be solved by introducing slack variables. The use of slack variables allow for some normal data points not included in the sphere description. Although, this can led to a number of anomalies lying within the sphere as well.

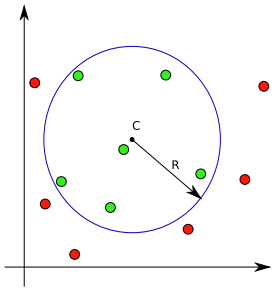


**Fig. 1.** The geometric representation of the sphere formulation

The task to minimize the volume of the sphere can be mathematically described as:

$$\begin{array}{ll} \underset{R\in\Re,\xi\in\Re^l,c\in\mathcal{F}}{min} & R^2+\frac{1}{\nu l}\sum_{i=1}^{l}\xi_i, \\ \text{subject to}: & ||\Phi(x_i)-c||\leq R^2+\xi_i \\ & \xi_i>0 \end{array} \tag{1}$$

The center $c$ of the hypersphere lies within the high-dimensional feature space $\mathcal{F}$. The non-negative slack variables $\xi_i$ allow for some data points to lie outside the hypersphere. The constant $\nu$ gives the trade-off between the two terms: volume of the sphere and the number of target objects rejected.

Since the center $c$ lies within the high-dimensional feature space, it is not possible to directly solve the primal problem (1) of the sphere formulation. The following dual problem where all the variables have low dimensions is solved instead:

$$\begin{array}{ll} \underset{\alpha \in \Re^l}{min} & \sum_{ij=1}^{l} \alpha_i \alpha_j k(x_i, x_j) - \sum_{i=1}^{l} \alpha_i k(x_i, x_i), \\ subject\,to: & \sum_{i=1}^{l} \alpha_i = 1, \\ & 0 \leq \alpha_i \leq \frac{1}{\nu l}. \end{array} \tag{2}$$

The mapping of datapoints to a high-dimensional feature space is defined by the kernel function $k(x_i, x_j)$, a generalization of the inner product in the feature space. Commonly used kernels are linear, sigmoid, polynomial, among others.

One of the most successful kernels used in the field of network traffic anomaly detection is the radial basis function (RBF), shown in following equation:

$$k(x_i, x_j) = e^{-\gamma(x_i - x_j)^2} \tag{3}$$

Where $\gamma = \frac{1}{\sigma^2}$ . Notice the parameter $\gamma$ gives the width, or spread, of the kernel function.

The classification between normal and anomalous traffic is done through the decision function, computed as follows:

$$f(x) = sgn(R^2 - \sum_{ij=1}^{l} \alpha_i \alpha_j k(x_i, x_j) + 2 \sum_{i=1}^{l} \alpha_i k(x_i, x_i) - k(x, x)) \tag{4}$$

The radius $R^2$ plays the role of a threshold, and it can be computed by equating the expression under the *sgn* to zero for any support vector.

### 2.2 Previous work on SVM for anomaly detection in intrusion detection

Different authors [3,4,5] have used support vector machines for novelty detection in the intrusion detection field. The work of Eskin et al. [3] is one of the first on the subject. They propose a geometrical framework to improve the performance of different kinds of unsupervised learning algorithms among which SVM is found. Laskov et al.[5] used the same geometrical framework presented by Eskin [3] and they provide a modification to SVM for anomaly detection which outperforms traditional variants. Both works use the KDD99 DARPA dataset for training and evaluating their approach.

The work of Li et al. [4] propose an improvement on SVM for novelty detection applied to the intrusion detection field. The idea is basically to extend Scholkopf's [11] hyperplane-to-origin approach. In their article, they assume that not only the origin lies in the second class but also that all data points close enough to the origin are to be considered as outliers or anomaly data points. For the evaluation process of their approach the authors use the 1999 DARPA dataset.

It seems clear that all these authors are aware of the limitations of the different SVM approaches for anomalies detection. As mentioned by Eskin [3], these algorithms will work well under the assumption that the number of normal traffic instances vastly outnumbers the number of anomalies. Moreover, in the experiments conducted, the authors assume that a high unbalance in class distribution is a common feature in network traffic and they have altered the original data sets to fit into this assumption. Unfortunately in practice, this assumption is not always true. There are many situations in which for specific periods of time, the presence of intrusions vastly exceeds the number of normal traffic instances. For instance, when a new vulnerability is discovered and it has been widely announced, it is possible to find attacks exploiting these vulnerability encompassing a extremely high percentage of the network traffic. Thus, it seems that anomalies in network traffic have a bursty behavior. This can be observed in the DARPA dataset, where the percentage of anomalous traffic found in some weeks is less than 0.5% but in some other weeks the percentage raise to 70%. However, this dataset may not be representative of the actual unbalanced in a production environment. The authors are unaware of a thorough study that confirms these claims.

Preliminary experiments conducted on the 1998 DARPA data set by the authors [13] confirms a poor performance for SVM for anomalies detection when data set does not present a highly unbalance class distribution.

On the other hand, the results obtained when SVM algorithms were trained using a highly unbalanced class distribution were similar to those reported by Eskin [3], as expected.

In a real traffic situation, it seems clear that it is not always possible to guarantee the required unbalance in class distribution for training sets, as needed by SVM approaches. A possible solution is to rely on experts for removing known attacks from the training set, until the desired unbalance is reached. This, however, would be an extremely expensive and tedious task. More interesting is the idea of using an autonomous labeling tool for removing known attacks.

## 3 Proposed approach: Autonomous labeling of normal traffic using SNORT.

For dealing with non-unbalanced class distribution situation an autonomous labeling approach is proposed. The idea is to use the attacks recognized by SNORT, a misuse signature-based IDS, to reduce the number of attacks in the training data set, and then use this reduced version of the data set to train a SVM for anomaly detection algorithm. The assumption is that after the attacks recognized by SNORT are removed from the training data set, the number of normal traffic instances will be sufficiently larger than the number of attacks. This way, class distribution becomes unbalanced or at least closer to the suggested unbalance.

SNORT [7] is a light and fast intrusion detection system developed by Martin Roesch in 1999. Over the past years, its popularity grew considerably, becoming a de-facto standard in the security network field. SNORT is composed by several fast pattern matching algorithms and a very complete and updated rule database. However, SNORT is far from being a complete solution to the intrusion problem. As any other misuse signature-based IDS, SNORT fails to recognize attacks which are not describe by a rule of its database. Another well known problem is that in many cases, SNORT can raise an extremely high false alarm rate, leading to production of different approaches for reducing SNORT false alarm [14].

The main hypothesis of this work is that although SNORT presents some drawbacks for classification, it still can be useful for labeling normal traffic, producing potentially better results of SVM for anomaly detection.

## 4 Experiments

A number of experiments were conducted in order to compare the behavior of the SNORT-based SVM anomaly detection approach, denoted as *SbSVM* with the standard SVM approach for anomaly detection.

### 4.1 Data set description

The experiments were conducted over five weeks of the 1998 DARPA data set [6], widely used for intrusion detection evaluation.

A total of six fields from a network traffic instance were selected for describing the input data: connection time, protocol type, source port, destination port, source IP address and destination IP address. Selected fields are represented according to Table 1 resulting a total of 14 attributes used for training SVM for anomaly detection alternatives.

**Table 1.** Features representation

| Feature | Size |
|---|---|
| Connection time | 3 |
| Protocol Type | 1 |
| Source port | 1 |
| Destination port | 1 |
| Source IP address | 4 |
| Destination IP address | 4 |

To improve SVM performance and to avoid possible numerical problems, the features are normalized between the interval [0,1] as suggested in [15].

### 4.2 Standard performance metrics for IDS evaluation

Standard performance metrics for IDS evaluation are used for comparing the different approaches discussed. These metrics correspond to *accuracy*, *attack detection* rate and *false alarm* rate

Accuracy is computed as the ratio between the number of correctly classified traffic instances and the total number of traffic instances. Detection rate is computed as the ratio between the number of correctly detected attacks and the total number of attacks. Finally, false alarm rate is computed as the ratio between the number of normal connections that are incorrectly classified as attacks and the total number of normal connections.

### 4.3 Standalone SNORT evaluation

In this section, the classification performance (in normal traffic and attacks) of a standalone SNORT is evaluated over the complete DARPA data set. From a total of thousands of rules in the SNORT rule-base, only 35 matched against the whole 5 weeks of the DARPA data set. Thus, for improving further computations the unmatching rules were removed from SNORT's rule database.

It is important to investigate the influence the size of SNORT's rule base has on the accuracy, detection rate and false alarm of SNORT against the DARPA data set. Thus, experiments were conducted using a SNORT classifier containing a rule-base of sizes 5, 10, 25, 30, 32 and 35, the total of the available rules. For statistical significance, for each rule-base size, a total of 10 subsets of that size were considered, and for each such subset, 24 repetitions of the experiments were conducted each on a randomly selected 0.5% subset of the DARPA data set.

The averaged results and the standard deviation are presented in Fig 2. It can be observed that as the size of the rule database increases, SNORT's accuracy and the attack detection rate increases as well. As expected the best results were achieved when SNORT uses the complete rule database. In that case the average accuracy obtained is 59%, the detection attack rate is 49%, and a very high false alarm rate of 29% is observed as well. In some cases, results present a significant variance. This behavior can be explained by the attacks distribution in the DARPA data set. The DARPA dataset contains a significative number of attacks of the type which can't be detected by SNORT. This situation was already reported in a previous work by Brugger et al. [16].

The obtained results seem to indicate SNORT performance on the DARPA data set is not very accurate. However, attack detection rate presented by SNORT for rule base sizes greater than 25 can still be useful for labeling a significative amount of normal traffic. Therefore, it is expected that in those cases *SbSVM* can bring class distribution closer to the unbalance required by SVM algorithms for anomaly detection.

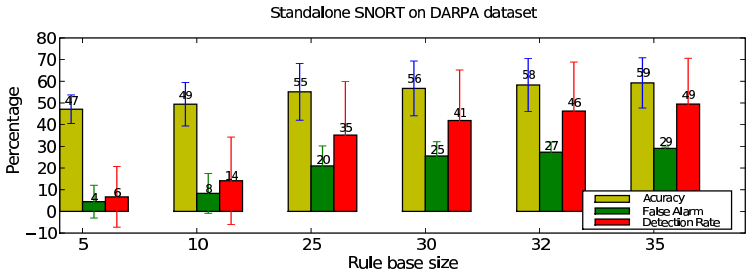


**Fig. 2.** Influence of the number of rules used by SNORT on the DARPA data set

### 4.4 Evaluation of the SNORT-based autonomous labeling for SVM anomaly detection

In order to evaluate the performance of the proposed *SbSVM* approach, the obtained results are compared with the ones computed using standard SVM.

The SVM implementation used for these experiments is an extension of the *libsvm* [17] that support the hypersphere formulation [18]. For these experiments, a RBF kernel with $\gamma = 8$ was selected and the chosen value for the penalty factor $\nu$ was 0.27. These values were obtained following the grid search procedure described in [15].

For training purposes, 1% subset of the DARPA data set was used, whereas another 0.5% subset was used for testing purposes, following standard ratios used in classification problems. This process was repeated 24 times for different randomly and uniformly selected subsets of DARPA.

The training process of the standard SVM approach uses the whole 1% including both normal and anomalous traffic. In the case of the *SbSVM* approach, instead, attacks recognized by SNORT are removed.

It seems interesting to investigate the influence of the rule database size used by SNORT on the SVM for anomaly detection approach. Thus, as in the previous experiment, rule database of different size were selected. Finally, to improve statistics estimators, each time a new 0.5% subset was selected for evaluation, the complete rule set was shuffle for a total of ten times.

The averages obtained are shown in Fig. 3. For comparison purposes obtained results by standalone SNORT (previously discussed in sub. 4.3) are also included.

The average accuracy achieved by standard SVM approach was 30%. This is just a minor difference compared to the 34% accuracy obtained when the training data set was labeled by SNORT with a rule database size of 5. The major differences can be observed when more than 25 rules are present at the SNORT database. In those cases, the average accuracy obtained is 63%, 68% and 72% for 25, 30 and 32 rules, respectively. Finally, when the training data set was labeled by the complete rule database, a 77% accuracy was achieved.

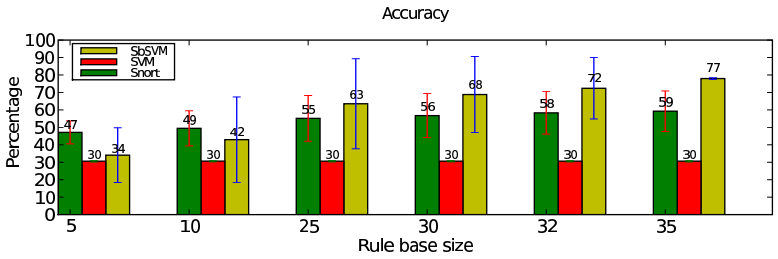


**Fig. 3.** Accuracy obtained by the *SbSVM* approach compared to standard SVM for anomalies detection and standalone SNORT.

In Fig. 4 results for attack detection rate are shown. The use of *SbSVM* approach shows significant improvement over the standard SVM. When SNORT uses a rule database size of 5, an average 13% attack detection rate is obtained, against a 6% detection rate of standard SVM.

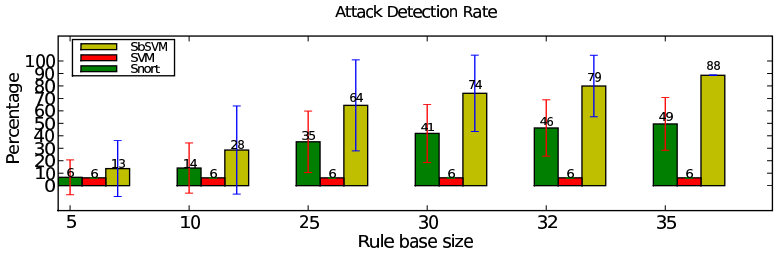


**Fig. 4.** Detection attack rate obtained by the *SbSVM* approach compared to standard SVM for anomalies detection and standalone SNORT.

Moreover, the attack detection rate increases as the rule database size increases, with a maximum attack detection rate of 88% obtained when the complete rule database was used.

Results for false alarm rate are shown in Fig. 5. As can be observed the false alarm rate obtained by the standard SVM variant was an average 40%. While in the cases where *SbSVM* was used, for almost any case, the average false alarm rate decreased (exception for 5 rules). When the complete rule base is used a false alarm rate of 35% is achieved.

Finally, when comparing *SbSVM* performance against results obtained by standalone SNORT (shown in Fig. 2), *SbSVM* shows better accuracy in all cases except when SNORT uses a rule base size of 5 and 10. More remarkable are the results obtained for attack detection rate. In every case, SbSVM shows major performance improvements over standalone SNORT. In the case when SNORT uses the complete rule base, attacks recognized by *SbSVM* are 80% more than

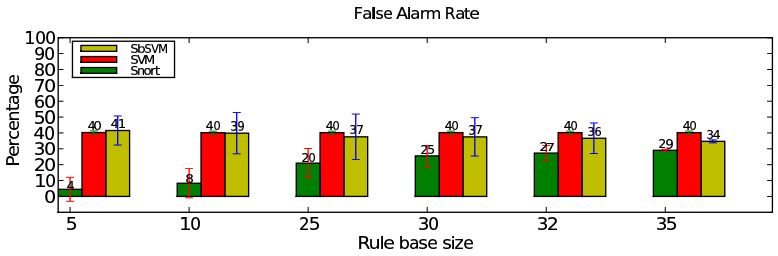


**Fig. 5.** False alarm rates obtained by the *SbSVM* approach compared to standard SVM for anomalies detection and standalone SNORT.

the ones recognized by standalone SNORT. For *SbSVM* a 34% of false alarm rate has been obtained instead of a 29% for standalone SNORT.

## 5 Conclusions

The performance obtained by standard SVM variant based on the hypersphere formulation on the 1998 DARPA dataset seems to confirm what has been already discussed in section 2.2. When a high number of attacks are included in the dataset, SVM algorithms for anomaly detection are not suitable for finding an accurate domain description. Thus, a highly unbalanced class distribution is needed in the dataset to achieve a proper performance.

The use of SNORT as an autonomous labeling tool appears to be a promising strategy to overcome this issue. Although in some cases SNORT exhibits a low attack detection rate, this situation has not prevented important performance improvements. For instance, accuracy is about of 50% better than the one obtained using standard SVM approach. Moreover, the attack detection are computed with the *SbSVM* approach shows an improvement of more than ten times compared with the one obtained with standard SVM.

The false alarm rate was also improved with proposed approach, although the improvements were not as significant as in the other performance metrics. The 34% false alarm rate obtained when the complete rule dataset is used by SNORT could be considered high for practical cases. This high false alarm rate value can be explained by the high false alarm rate obtained by standalone SNORT on the DARPA dataset.

The obtained results have shown the autonomous labeling approach using SNORT has improved not only SVM algorithms for anomaly detection but also the standalone SNORT.

The performance of the *SbSVM* approach in real traffic situations still remains unknown. Consequentely, experiments will be carried out to overcome this issue in the future.

## References


1. B. Mukherjee, L. T. Heberline, and K. Levitt, "Network instruction detection.", *IEEE Network*, vol. 8, pp. 26–41, 1994.
2. C. Catania and C. García Garino, "Reconocimiento de patrones en el tráfico de red basado en algoritmos genéticos", *Inteligencia Artificial, Revista Iberoamericana de IA*, vol. 12, no. 37, pp. 65–75, 2008.
3. Eleazar Eskin, Andrew Arnold, Michael Prerau, Leonid Portnoy, and Sal Stolfo, "A geometric framework for unsupervised anomaly detection: Detecting intrusions in unlabeled data", in *Applications of Data Mining in Computer Security*. 2002, Kluwer.
4. Kun-Lun Li, Hou-Kuan Huang, Sheng-Feng Tian, and Wei Xu, "Improving one-class svm for anomaly detection", Nov. 2003, vol. 5, pp. 3077–3081 Vol.5.
5. Pavel Laskov, Christin Schafer, and Igor Kotenko, "Intrusion detection in unlabeled data with quarter-sphere support vector machines", in *In Proc. DIMVA*, 2004, pp. 71–82.
6. R. Lippmann, J. W. Fried, D. J. Korba, and K Das, "The 1999 darpa off-line intrusion detection evaluation", *Computer Networks*, vol. 34, pp. 579–595, 2000.
7. Martin Roesch, "Snort - lightweight intrusion detection for networks", in *LISA '99: Proceedings of the 13th USENIX conference on System administration*, Berkeley, CA, USA, 1999, pp. 229–238, USENIX Association.
8. Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik, "A training algorithm for optimal margin classifiers", in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*. 1992, pp. 144–152, ACM Press.
9. Corinna Cortes and Vladimir Vapnik, "Support vector networks", pp. 273–297, 1995.
10. David M. J. Tax and Robert P. W. Duin, "Data domain description using support vectors", in *Proceedings of the European Symposium on Artificial Neural Networks*, 1999, pp. 251–256.
11. Bernhard Scholkopf, John C. Platt Z, John Shawe taylor Y, Alex J. Smola X, and Robert C. Williamson X, "Estimating the support of a high-dimensional distribution", *Neural Computation*, vol. 13, pp. 1443–1471, 2001.
12. Colin Campbell, "Kernel methods: A survey of current techniques", *Neurocomputing*, vol. 48, pp. 63–84, 2000.
13. Carlos Catania, "Enfoques para detección de intrusos en el tráfico de red basados en máquinas de vectores soporte.", Tech. Rep., ITIC - UNCuyo, 2009.
14. Gina C. Tjhai, Maria Papadaki, Steven M. Furnell, , and Nathan L. Clarke and, "The problem of false alarms: Evaluation with snort and DARPA 1999 dataset", in *Trust, Privacy and Security in Digital Business*. 2008, vol. 5185 of *Lecture Notes in Computer Science*, pp. 139–150, Springer Berlin / Heidelberg.
15. Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, "A practical guide to support vector classification", URL http://www.csie.ntu.edu.tw/ cjlin/papers/guide/guide.pdf, 2008.
16. S Terry Brugger and Jedadiah Chow, "An assessment of the darpa ids evaluation dataset using snort", 2005.
17. Chih-Chung Chang and Chih-Jen Lin, *LIBSVM: a library for support vector machines*, 2001, Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.
18. Vicenzo Russo, "Libsvm plus", URL http://neminis.org/software/libsvm-plus/, 2008.