# Intrusion Detection System Based on Multi-class SVM

Hansung Lee, Jiyoung Song, and Daihee Park

Dept. of computer & Information Science, Korea Univ., Korea
{mohan, songjy, dhpark}@korea.ac.kr

**Abstract.** In this paper, we propose a new intrusion detection system: MMIDS (Multi-step Multi-class Intrusion Detection System), which alleviates some drawbacks associated with misuse detection and anomaly detection. The MMIDS consists of a hierarchical structure of one-class SVM, novel multi-class SVM, and incremental clustering algorithm: Fuzzy-ART. It is able to detect novel attacks, to give detail informations of attack types, to provide economic system maintenance, and to provide incremental update and extension with a system.

## 1 Introduction

Recently threat elements and criminal behaviors against computer and information resources on networks appears to increase and communize rapidly to the extent harming on household computers. Threat elements against computer resources are expanded to a variety of types of attack ranging from simple intrusion by use of scripts to misuse of computers by use of malware equipped with multiple functions, and the scale of damage also increases rapidly [1]. Hence, it is necessary to develop more effective intrusion detection algorithms in order to detect malicious behaviors and intrusions being expanded to a variety of range.

Intrusion detection is the art of detecting unauthorized, inappropriate, or anomalous activity on computer systems. There are two major paradigms for intrusion detection methods according to general strategy for detection: misuse detection and anomaly detection. The misuse detection model establishes a rule base by way of a detail analysis on the attack type, and performs the detection on the basis of it. This model has an intrinsic disadvantage that the rule base should be manually updated for new types of attack in order to deal with them. The anomaly detection model detects attacks by determining data deviated at a great extent from the profile including normal behavior as abnormal behavior. This model is useful to detect new types of attack, but has an unavoidable limit that it can not take a proper action against the intrusion because it lacks detail information about the detected type of attack [2].

According an investigation on latest research literatures, there are many ongoing attempts to apply data mining and machine learning techniques to the intrusion detection systems in order to design more intelligent intrusion detection model. Recently, the support vector learning method, featuring superior

D. Ślęzak et al. (Eds.): RSFDGrC 2005, LNAI 3642, pp. 511–519, 2005.

performance at some issues such as pattern classification and function approximation, has grown up as a viable tool in the area of intrusion detection systems. The intrusion detection model based on SVM (support vector machine) is mainly classified into three types. The first type [3] divides data into normal data and attack data using characteristics of the binary classifier SVM, and the second type [4] implements an anomaly detection model using one-class SVM. All of both methods described above, however, can not offer additional information about detected type of attack, only but distinguish normal data and abnormal data. Finally, the third type [5] establishes multi-class SVM at the form of combining many binary classifiers, namely, SVMs and divides data into normal data and four types of attack data. Since this type has many problems including that system performance of which is completely dependent on the quality of training data and training time of which take a long time, it can not be used as practical intelligent intrusion detection system although it is more advanced one in comparison with other two methods.

In this paper, we propose a novel intrusion detection model, which keeps advantages of existing misuse detection model and anomaly detection model and resolves their problems. This new intrusion detection system, named to MMIDS (Multi-step Multi-class Intrusion Detection System), was designed to satisfy all the following requirements by combining one-class SVM, proposed novel multi-class SVM and Fuzzy-ART that is a incremental clustering algorithm, hierarchically: 1)Fast detection of new types of attack unknown to the system; 2) Provision of detail information about the detected types of attack; 3) cost-effective maintenance due to fast and efficient learning and update; 4) incrementality and scalability of system.

The organization of this paper is as follows. Section 2 explains the new-proposed multi-class SVM, and Section 3 describes the proposed intrusion detection model: MMIDS. The experimental results are given in Section 4, and conclusions are made in Section 5.

## 2 Multi-class SVM Based on One-Class SVM

Recently, the support vector learning method has grown up as a viable tool in the area of intelligent systems. It shows an excellent performance at the pattern classification and function approximation by ensuring global optimum for a given problems. However, it is difficult to apply SVM at real world applications, most of which are as to the multi-class classification, because it has a structural limit of the binary classifier. How to effectively extend it for multi-class classification is still an ongoing research issue. Currently, there are three major types of approaches for multi-class SVM: one-against-all, one-against-one, and DAGSVM [6].

As for the intrusion detection, the volume of data necessary for the training varies depending on each type of attack. Hence, the training result for a type of attack may be affected by other types of attack data owing to the unbalanced size of training data. In addition, it can not be said that current training data

represents for the whole class, because new types of attack, included in one attack class, are increasingly emerged. So, the binary classifier SVM may be subject to misclassification for novel attack data by creating decision boundary including unobserved area. It is preferable, therefore, to select the decision boundary using one-class SVM [8] that expresses only the corresponding class independently. In this paper, we propose a multi-class SVM that can classify various types of attack on the basis of one-class SVM.

Given $K$-data set of $N_k$ patterns in $d$-dimensional input space, $D_k = \{x_i^k \in R^d | i = 1, \ldots, N_k\}; k = 1, \ldots, K$, multi-class SVM to classify each class is defined as a problem to obtain a sphere that minimizes the volume with it including the training data, and it is formalized through the following optimization problem.

$$\begin{aligned} min \quad & L_0(R_k^2, a_k, \xi_k) = R_k^2 + C \sum_{i=1}^{N_k} \xi_i^k \\ s.t. \quad & \|x_i^k - a_k\|^2 \leq R_k^2 + \xi_i^k, \xi_i^k \geq 0, \forall i. \end{aligned} \tag{1}$$

Where, $a_k$ is the center of the sphere that expresses $k$-th class, $R_k^2$ is the square value of sphere radius, $\xi_i^k$ is the penalty term that shows how far $i$-th training data $x_i^k$ pertained to $k$-th class is deviated from the sphere, and $C$ is the trade-off constant.

By introducing a Lagranger multiplier for each inequality condition, we obtain the following Lagrange function:

$$\begin{aligned} L(R_k^2, a_k, \xi_k, \alpha_k, \eta_k) = R_k^2 + C \sum_{i=1}^{N_k} \xi_i^k \\ + \sum_{i=1}^{N_k} \alpha_i^k \left[ (x_i^k - a_k)^T (x_i^k - a_k) - R_k^2 - \xi_i^k \right] \\ - \sum_{i=1}^{N_k} \eta_i^k \xi_i^k \end{aligned} \tag{2}$$

where $\alpha_i^k \geq 0$, $\eta_i^k \geq 0$, $\forall i$.

From the saddle point condition, the equation (2) has to be minimized with respect to $R_k^2$, $a_k$ and $\xi_i^k$ and maximized with respect to $\alpha_k$ and $\eta_k$ [7,8]. The optimal solution of (1) should satisfy the following:

$$\begin{aligned} & \frac{\partial L}{\partial R_k^2} = 0 : \sum_{i=1} N_k a_i^k = 1. \\ & \frac{\partial L}{\partial \xi_i^k} = 0 : C - \alpha_i^k - \eta_i^k = 0 \quad \therefore \alpha_i^k \in [0, C], \forall i. \\ & \frac{\partial L}{\partial a_k} = 0 : a_k = \sum_{i=1} N_k \alpha_i^k x_i^k. \end{aligned} \tag{3}$$

With substitution of the above into Lagrange function $L$, we obtain the following dual problem:

$$min_{\alpha_k} \quad \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k < x_i^k, x_j^k > - \sum_{i=1}^{N_k}\alpha_i^k < x_i^k, x_i^k > \tag{4}$$

$$s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1, \alpha_i^k \in [0, C], \forall i.$$

Since, each class can express more complex own decision boundary in respective feature space $F$ [7,8], training of the system is done with solving the following QP problem by considering independency of the feature space to which each class is mapped.

$$min_{\alpha_k} \quad \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k K_k(x_i^k, x_j^k) - \sum_{i=1}^{N_k}\alpha_i^k K_k(x_i^k, x_i^k) \tag{5}$$

$$s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1, \alpha_i^k \in [0, C], \forall i.$$

When the gaussian function is chosen for the kernel, we always have $K(x, x) = 1$ for each $x \in R^d$. Thus, the above problem can be further simplified as follows:

$$min_{\alpha_k} \quad \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k K_k(x_i^k, x_j^k) \tag{6}$$

$$s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1, \alpha_i^k \in [0, C], \forall i.$$

Note that in this case, the decision function of each class can be summarized as follows

$$f_k(x) = R_k^2 - \big(1 - 2\sum_{i=1}^{N_k}\alpha_i^k K_k(x_i^k, x) + \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k K_k(x_i^k, x_j^k)\big) \geq 0 \tag{7}$$

Since the output $f_k(x)$ of one-class SVM that is defined on different feature space means absolute distance between corresponding data and decision boundary in each feature space, it is not recommended to determine pertained class by comparing absolute distances on different feature spaces. Accordingly, we calculate the relative distance $\hat{f}_k(x) = f_k(x)/R_k$ by dividing the absolute distance $f_k(x)$ on the feature by the radius $R_k$ of the sphere that is defined on the feature space, and decide the class having maximum relative distance as one to which the input data $x$ is pertained.

$$Class\ \ of\ \ x \equiv arg\ max_{k=1,\ldots,k}\hat{f}_k(x) \tag{8}$$
$$where\ \ \hat{f}_k(x) = f_k(x)/R_k$$

**Remark:** To show its effectiveness, we compare the proposed method with other methods with respect to complexity analysis. Let's assume $k$ classes and $n$ training data per class in order to compare the multi-class SVM that we propose and conventional methodologies. For one-against-all, total number of data for the training is $nk^2$ because the number of SVMs to be trained is $k$ and each SVM has to train data amounting to $nk$. For one-against-one and DAG, the number of SVMs to be trained is $k(k-1)/2$ and $2n$ data contributes to the training of each SVM. Total number of data to be trained is, therefore, $nk^2 - nk$. For the proposed method, with total number of data involved in the training being $nk$ because $k$ SVMs are trained and $n$ data per SVM is involved, the proposed algorithm delivers a faster training speed. If one new class is added into the multi-class SVM that has already been trained, $k+1$ SVMs should be trained newly for one-against-all and $k$ SVMs should be trained again for one-against-one, respectively. In the case of DAG, not only $k$ SVMs should be trained again but also additional cost for reconstructing the graph is required. Whereas, the proposed algorithm is very cost-effective in terms of system reconstruction because it trains only added SVM. The Complexity Analysis of conventional Multi-Class SVMs and proposed method are summarized in Table 1.

**Table 1.** Complexity Analysis of Multi-Class SVMs

| | The number of training SVMs | The number of training data per SVM | The number of training SVMs when a class is added | The number of test SVMs |
|---|---|---|---|---|
| one-against-all | $k$ | $nk$ | $k+1$ | $k$ |
| one-against-one | $k(k-1)/2$ | $2n$ | $k$ | $k(k-1)/2$ |
| DAG | $k(k-1)/2$ | $2n$ | $k + DAG$ | $k$ |
| Proposed Method | $k$ | $n$ | 1 | $k$ |

## 3 MMIDS (Multi-step Multi-class Intrusion Detection System)

In this section, we propose a new intrusion detection model, which keeps advantages of existing misuse detection model and anomaly detection model as they are and resolves their problems. The new intrusion detection system, named to MMIDS, is designed to satisfy all the following requirements: 1) Fast detection of new types of attack unknown to the system; 2) Provision of detail information about the detected types of attack; 3) cost-effective maintenance due to fast and efficient learning and update; 4) incrementality and scalability of system.

As shown in Figure 1, MMIDS consists of three main components; one-class SVM that distiguishes normal data and attack data; multi-class SVM classifies the attack data into one of DOS (denial of service), R2L (remote to local), U2R (user to root) and Probing attacks; Fuzzy-ART that carries out detail clustering

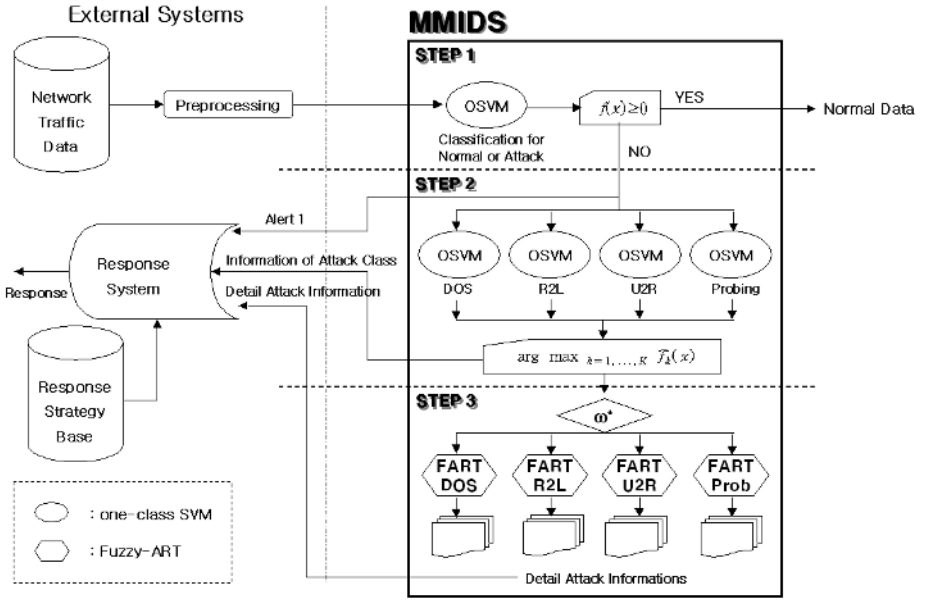


**Fig. 1.** The Architecture of MMIDS

for each type of attack. There are three phases in the procedure of intrusion detection for the test data. We take a look at them closely here.

**Phase 1:** One-class SVM that has been trained by normal data classifies normal data and attack data at the first stage. The training process requires only normal data without need to prepare attack data, ensuring a faster training speed. The one-class SVM, as an anomaly detection model, detects novel attack without additional process for normal data on the system operation. When attack data is detected, it generates first alarm to the intrusion response system and steps into the phase 2.

**Phase 2:** The multi-class SVM classifies the attack data detected at phase 1 into one of DOS, U2R, R2L and Probing, and provides the intrusion response system with additional information about the type of attack. For the update of the novel attack data to the system, only classifier of the corresponding class rather than whole system is retrained. It allows the maintenance cost on the on-site system operation to be reduced.

**Phase 3:** If more detail information about incoming attack is required, Fuzzy-ART performes clustering for each type of attack. Detail data for each attack is not much enough to train classifiers such as SVM, and it is difficult to predict its type as well. Thus it is recommended to carry out classification by attack using the unsupervised learning, namely, clustering.

## 4 Experimental Results

To evaluate the effectiveness of the intrusion detection system, MMIDS, KDD CUP 1999 data [10] were used for the experiments. In order to make accurate analysis on the experiment result, we used only Corrected-labeled data set among KDD CUP 99 data which was collected through the simulation on the U.S. military network by 1998 DARPA Intrusion Detection Evaluation Program, aiming at obtaining the benchmark data set in the field of intrusion detection. The used data amounting to 311,029 consists of 9 symbolic attributes and 32 numeric attributes. The data is mainly divided into four types of attack: DOS, R2L, U2R and probing, and subsidiary types of attack by main attack amounts to 37 including 6 Probing attacks such as ipsweep and saint.

### 4.1 Comparison with Other Intrusion Detection Algorithms

Because many research results of intrusion detection have been reported recently, we compare our performance with other Intrusion Detection Algorithms. Table 2 shows classification capability of each research for normal data and 4 types of attack. The multi-class classifier proposed in this paper has superior classification capability to conventional researches as a whole, as shown in Table 2. It should be observed that all research results show considerable inferior performance only at the classification capability as to R2L and U2R. As shown in Table 2, our method can provide superior performance in separating these two patterns. In particular in the comparison between this experiment and Ambwani [5] that used one-against-one multi-class SVM targeting at same data as ours, the experiment results of this methodology shows much enhanced performance. It can be explained as follows;

Note that R2L and U2R are host-based attacks which exploit vulnerabilities of operating systems, not of network protocol. Therefore, these are very similar to the "normal" data in the KDD CUP 99 data collected from network packets. And depending on type of attack, the size of data available for the training is quite different. The size of U2R or R2L attack data, for example, is smaller than that of other types of attack. Hence at the conventional multi-class SVM that trains decision boundary for each class after mapping all classes into the same feature space, the class, size of which is relatively larger, may deliver more impact on the training result than others. It can be said, therefore, that the strategy of multi-class SVM employed in this paper, that is to enhance classification performance

**Table 2.** Comparison with Other Intrusion Detection Algorithms

| | Bernhard [10] | W. Lee [11] | Y. Liu [12] | Kayacik [13] | Ambwani [5] | MMIDS |
|---|---|---|---|---|---|---|
| Normal | 99.5% | - | - | 95.4% | 99.6% | **96.74%** |
| DOS | 97.1% | 79.9% | 56% | 95.1% | 96.8% | **98.24%** |
| **R2L** | **8.4%** | **60.0%** | **78%** | **9.9%** | **5.3%** | **35.00%** |
| **U2R** | **13.2%** | **75.0%** | **66%** | **10.0%** | **4.2%** | **85.23%** |
| Probing | 83.3% | 97.0% | 44% | 64.3% | 75.0% | **98.27%** |

for own-class and minimizes the effect from unbalanced size of training data (see Eq. 6), is turned out to be very efficient.

### 4.2 Clustering of Attack Data

In addition, we performed the clustering for each of 4 types of attack using Fuzzy-ART in order to verify detailed separation capability corresponding to phase 3 of MMIDS. Table 3 shows the results from the Probing attack. According to the experimental results of Table 3, detailed separation capability of MMIDS is relatively good.

**Table 3.** Experimental Results of Probing Attack

| Attacks | Detection Ratio | Attacks | Detection Ratio |
|---|---|---|---|
| ipsweep | 95.0% | satan | 91.7% |
| saint | 17.2% | mscan | 88.2% |
| portsweep | 96.0% | nmap | 20.0% |

## 5 Conclusions

In this paper, we proposed a new intrusion detection model so that it keeps advantages of existing misuse detection model and anomaly detection model and resolves their problems. This novel intrusion detection system, named to MMIDS, was designed to satisfy all the following requirements by combining one-class SVM, proposed multi-class SVM and Fuzzy-ART that is an incremental clustering algorithm, hierarchically: 1) Fast detection of new types of attack unknown to the system; 2) Provision of detail information about the detected types of attack; 3)cost-effective maintenance due to fast and efficient learning and update; 4) incrementality and scalability of system.

Comprehensive researches that consider organic relationship between two systems: IDS and response system, including the method of utilizing detail information about attacks detected by the intrusion detection system for establishing policy of intrusion detection system, may be required for future works.

## References

1. E. Skoudis and L. Zeltser: Malware - Fighting Malicious Code, Prentice Hall, 2004.
2. S. Noel, D. Wijesekera, and C. Youman: 'Modern Intrusion Detection, Data Mining, and Degrees of Attack Guilt', in Applications of Data Mining in Computer Security, Kluwer Academic Publisher, pp. 1-31, 2002.
3. W.H. Chen, S.H. Hsu, and H.P. Shen: 'Application of SVM and ANN for intrusion detection', Computers & Operations Research, ELSEVIER, Vol. 32, Issue 10, pp. 2617-2634, 2005.

4. K.L. Li, H.K. Huang, S.F. Tian, and W. Xu: 'Improving one-class SVM for anomaly detection', International Conference on Machine Learning and Cybernetics, Vol. 5, pp. 3077-3081, 2003.
5. T. Ambwani: 'Multi class support vector machine implementation to intrusion detection', Proceedings of the International Joint Conference on Neural Networks, Vol. 3, pp. 2300-2305, 2003.
6. C.W. Hsu and C.J. Lin.: 'A comparison of methods for multi-class support vector machines', IEEE Transactions on Neural Networks, Vol. 13, pp. 415-425, 2002.
7. N. Cristianini and J. Shawe-Taylor: 'An introduction to support vector machines and other kernel-based learning methods', Cambridge University PRESS, pp. 93-124, 2000.
8. D.M.J. Tax and R.P.W. Duin: 'Uniform Object Generation for Optimizing One-class Classifiers', Journal of Machine Learning Research, Vol. 2, Issue 2, pp. 155-173, 2001.
9. J. Huang, M. Georgiopoulos, and GL Heileman: 'Fuzzy ART properties', Neural Networks, Vol. 8, No. 2, pp. 203-213, 1995.
10. Results of the KDD'99 Classifier Learning Contest, Available in http://www-cse.ucsd.edu/users/elkan/clresults.html
11. W. Lee, S.J. Stolfo, and K.W. Mok: 'A data mining framework for building intrusion detection models', Proceedings of the 1999 IEEE Symposium on Security and Privacy, pp. 120-132, 1999.
12. Y. Liu, K. Chen, X. Liao, and W. Zhang: 'A Genetic Clustering Method for Intrusion Detection', Pattern Recognition, Vol. 37, Issue 5, pp. 927-942. 2004.
13. H.G. Kayacik, A.N. Zincir-Heywood, and M.I. Heywood: 'On the capability of an SOM based intrusion detection system', Proceedings of the International Joint Conference on Neural Networks, Vol. 3, pp. 1808-1813, 2003.