Network Intrusion Detection System Based on Incremental Support Vector Machine

Haiyi Zhang, Yang Yi, and Jiansheng Wu

Abstract. Based on simple incremental SVM, we proposed an improved incremental SVM algorithm (ISVM), and combined it into a kernel function U-RBF and applied it into network intrusion detection. The simulation results show that the improved kernel function U-RBF has played some role in saving training time and test time. The ISVM has eased the oscillation phenomenon in the process of the learning to some extent, and the stability of ISVM is relatively good.

1 Introduction

Intrusion detection, a proactive real-time security protection technology, has got more and more attention. It is a network security technology that is used to detect acts, which damage or attempt to damage the confidentiality, integrity or availability of the system or network. It can effectively make up the shortages of firewalls, data encryption, authentication and other static defense strategies and can carry out a full range of network security protection by combining with other network security products as well.

Research on the incremental SVM learning algorithm in the intrusion detection system is of significance. Firstly, it is difficult for traditional security and defense strategies to meet the ever-changing needs of network security and short of static protection technologies. Secondly, current intrusion detection methods are mostly non-incremental learning algorithm. As the accumulation of the new incremental samples, the training time expenses will continue to increase. Thirdly, incremental learning can rapidly learn from the new incremental samples to modify the existing model. Time consumption is relatively small. Finally, compared with non-incremental algorithms, incremental learning algorithms are in a relatively small number of studies, especially the incremental SVM algorithm.

Haiyi Zhang
Jodrey School of Computer Science, Acadia University, Wolfville, Noca Scotia, Canada B4P 2R6
e-mail: haiyi.zhang@acadiau.ca

Yang Yi · Jiansheng Wu
Sun Yat-sen University, China

M. Ali et al. (Eds.): *Contemporary Challenges & Solutions in Applied AI*, SCI 489, pp. 91–96.
DOI: 10.1007/978-3-319-00651-2_13

Due to the oscillation phenomenon that simple incremental SVM will lead to in the follow-up learning process, we propose an improved incremental SVM algorithm ISVM. In order to reduce the number of samples and shorten the training time, the algorithm introduces sample selection process, and applies it to network intrusion detection. Compared to other algorithms, the test results indicate that the ISVM algorithm eases the oscillation phenomenon to some extent in the incremental learning process, and also saves training and prediction time.

2 Incremental Support Vector Machine

Support vector machines are a new type of learning machine based on statistical learning theory and structural risk minimization principle. The basic idea is to make a nonlinear mapping from the input space to high dimensional space. Then construct a classification hyper plane that separates the training data by a maximal margin.

The standard SVM requires that all samples be trained at the same time. If new samples are added, the SVM needs to be retrained to find a new optimal classification hyper plane, so the tolerance to noise for the SVM is not high. On the other hand, it is very difficult to obtain a complete training set in the beginning if lack of initial samples. The accuracy of the learning machine will be affected, so we hope the learning machine can continuously improve the learning accuracy by using the priori knowledge when new samples are added in.

Based on the above issues, there are a number of researchers working on the SVM with incremental learning function. Syed et al [1,4] first proposed the incremental learning algorithm of SVM. It divided the training set into N subsets, after training on a subset, the algorithm only retained the support vectors and discarded the other samples, and then added them into next subset to form a new training subset and trained on the new training subset.

3 Incremental Support Vector Machine Based on Reserved Set

3.1 Kernel Function U-RBF (Unitizing RBF)

We propose a new kernel function, U-RBF, according to the consideration unitizing the record. The U-RBF adds the mean value of the feature attributes and the mean square deviation to the RBF, which will cut down the training time.

3.2 The KKT Conditions and the Distribution of the Samples

In order to obtain the decision function when training the SVM, we have to solve the quadratic programming problem, where we need to use the optimality conditions of the optimization problem, Karush-Kuhn-Tucker (KKT) conditions.

The solving of SVM boils down to the solving of the convex quadratic programming problem, where the KKT conditions are defined as:

$$y_i f(x_i) - 1 \begin{cases} \geq 0 & \alpha_i = 0 \\ = 0 & 0 < \alpha_i < C \\ \leq 0 & \alpha_i = C \end{cases} \tag{1}$$

The KKT conditions (1) are equivalent to (2):

$$\begin{cases} \alpha_i = 0 \Rightarrow f(x_i) \geq 1 \quad or \quad f(x_i) \leq -1 \\ 0 < \alpha_i < C \Rightarrow f(x_i) = 1 \quad or \quad f(x_i) = -1 \\ \alpha_i = C \Rightarrow -1 \leq f(x_i) \leq 1 \end{cases} \tag{2}$$

We can get the classifier from the process of training the samples, where α_i is the Lagrange multiplier that the sample corresponds to. According to (2), we know that the samples that meet the requirement of the KKT conditions are the support vectors.

3.3 Reserved Set Strategy

In the simple incremental SVM, when the new incremental samples come, first, check whether all the samples meet the KKT conditions. The follow-up learning will lead to oscillation phenomenon if the initial samples are insufficient. This paper presents a reserved set strategy that will retain. In order to ease the oscillation phenomenon in the follow-up learning process, select samples from the reserved set to combine them with those samples that are contrary to the KKT conditions in the new incremental sample set and the original support vector set to form a new training set, according to the weight. After incremental learning, update the reserved set and the weights of the samples.

There are two key issues: one is how to select samples to construct the reserved set; the other is how to empower the value for each sample.

4 Simulated Experiments

4.1 Experiments Description

We take the benchmark KDD-CUP99 with nearly 5 million network records, as the dataset of the experiments. Each record is a grouped sequence that starts and terminates within the required timeframe when in line with the stated protocols. It also has a class identifier, which denotes either normal class or some specific attack class. There are 22 attack classes divided into the four categories.

In this paper, the detection rate, false alarm rate and correlation coefficient are used as the evaluation indicators for the intrusion detection. The purpose of the

incremental SVM proposed in this paper is not only to enhance the intrusion detection rate and reduce false alarm rate, but also to reduce the training time and the forecast time as much as possible. So, the training time and the forecast time are adopted as the evaluation indicators as well.

The simulation experiments are divided into two parts. In the first part we mainly verify the effectiveness of the improved kernel function U-RBF by comparing the effectiveness of U-RBF with RBF and POLY. I in the second part we mainly test the detection performance of the RS-ISVM by comparing it with the simple incremental SVM and the peer-to-peer incremental SVM.

4.2 Experiments of RS-ISVM

Our experiments are performed to verify the effectiveness of the improved SVM. First, the dataset is randomly divided into two subsets, each contains both normal and abnormal class, one is the source of the training data, and the other is the source of test data. Secondly, select 9 data sets at randomly, named I1 to I9, from the training subset as the incremental training set randomly, each set contains 300 normal samples and 300 abnormal samples, and any two training samples sets are not mixed. Thirdly, select the normal records and attack records, with the number of normal records are equal to the attack records, as the test set.

Table 1 The performance of the RS-ISVM using different parameters

Parameters		Training set (cc)		
η_1	η_2	I_1 (cc)	I_2 (cc)	I_3 (cc)
0.1	10	0.767	0.759	0.768
0.1	20	0.767	0.798	0.814
0.2	10	0.767	0.765	0.769
0.2	20	0.767	0.799	0.815
0.3	10	0.767	0.766	0.771
0.3	**20**	**0.767**	**0.806**	**0.821**
0.4	10	0.767	0.749	0.737
0.4	20	0.767	0.029	0.785
0.5	10	0.767	0.784	0.766
0.5	20	0.767	0.103	0.753

Because the strategy of selecting samples involves the scale factors η_1 and η_2, which need to be set, we choose the best combination of η_1 and η_2 by assessing algorithm performance with combinations of different parameters. Different scale factors η_1 and η_2 have been chosen for simulation and the comparison results of correlation coefficients are listed in Table 1, where I_1~I_3 are the incremental training subsets and cc denotes the correlation coefficient.

The best combination of the parameters η_1 and η_2 has been marked in bold. First, train the SVM on the sample set I_1, and then the incremental SVM will be updated when the new incremental sample sets I_2~I_9 arrive. This paper will compare the RS-ISVM with the simple incremental SVM (Simple-ISVM) [2,5] and the peer-to-peer incremental SVM (KKT-ISVM) [3,6] with the detection rate, false alarm rat, correlation coefficient, training time and forecast time. Since all the three methods are based on the C-SVM, all the three algorithms will use the RBF as the kernel functions. Moreover, RS-ISVM also uses U-RBF.

Table 2 Comparison of the training time and test time

	RS-ISVM(U)		RS-ISVM		Simple-ISVM		KKT-ISVM	
	TrD(s)	TeD(s)	TrD(s)	TeD(s)	TrD(s)	TeD(s)	TrD(s)	TeD(s)
I1	0.438	1.469	1.094	4.656	1.016	5.64	1.105	5.688
I2	0.718	1.985	2.078	6.218	3.844	14.813	7.547	16.203
I3	1.015	2.719	3.156	9.266	5.625	20.562	16.0	23.984
I4	1.328	3.391	5.797	11.719	9.093	23.922	30.437	29.656
I5	1.719	5.546	6.781	13.641	18.86	26.656	38.297	34.297
I6	0.984	1.00	8.156	15.532	18.047	35.047	48.688	38.547
I7	0.438	1.105	8.609	16.938	22.672	42.094	71.64	44.563
I8	0.453	1.11	10.485	18.969	33.094	47.969	83.203	50.016
I9	0.553	1.11	13.812	21.328	28.14	43.813	96.266	53.047

The comparison results of the training time and test time are listed in Table 2, where TrD and TeD represent the training time and test time respectively. As shown in Table 2, KKT-ISVM needs the most training time, because it has to do cross judging and more training. The training time of Simple-ISVM and RS-ISVM is in the acceptable range. However, RS-ISVM(U) has obvious advantages in training time, because it is clearly much shorter than RS-ISVM.The above results show that the improved kernel function U-RBF plays some role in saving the training and test time. On the other hand, compared to the Simple-ISVM and KKT-ISVM, the changes that RS-ISVM do to the original classifier are more reliable, and it will not cause large fluctuations in detection performance to the classifier. Moreover, with

cumulative incremental training on the new sample set, RS-ISVM continuously improves the detection performance.

5 Conclusions

We have proposed a strategy that is based on the reserved set, which retains those non-support vectors that are most likely to become support vectors to ease the oscillation phenomenon in the process of the incremental learning. At the same time, a concentric circle method has been proposed to select the samples to construct the reserved set. Lastly, an incremental SVM algorithm RS-ISVM that is based on the reserved set bas been proposed.

References

[1] Wang, X.D., Zheng, C.Y., Wu, C.M., Zhang, H.D.: New algorithm for SVM-based incremental learning. Computer Applications 10(26), 2440–2443 (2006)
[2] Laskov, P., Gehl, C., Kruger, S., Muller, K.: Incremental support vector learning: Analysis, Implementation and application. Journal of Machine Learning Research 7, 1909–1936 (2006)
[3] Shilton, A., Palamiswami, M., Ralph, D., Tsoi, A.: Incremental training of support vector machines. IEEE Transactions on Neural Networks 16, 114–131 (2005)
[4] Cheng, S., Shih, F.: An improved incremental training algorithm for support vector machines using active query. Pattern Recognition 40, 964–971 (2007)
[5] Liang, Z.Z., Li, Y.F.: Incremental support vector machine learning in the primal and applications. Neurocomputing (February 20, 2009)
[6] Deng, N.Y., Tian, Y.J.: A new method of data mining – support vector machines. Science Press (2004)