CHEMICAL ENGINEERING TRANSACTIONS
VOL. 71, 2018
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Xiantang Zhang, Songrong Qian, Jianmin Xu

ISBN 978-88-95608-68-6; **ISSN** 2283-9216
DOI: 10.3303/CET1871190

Information Security of Chemical Substation Based on ISVM

Hongyan Shen*, Fang Cheng, Xiaoying Lu

College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China
shy0907@163.com

This paper proposes an ISVM (Improved Support Vector Machine) based cybersecurity attack detector used to improve defense against industrial control network attack and guarantee a safe and stable operation of chemical power systems, which abstracts the cybersecurity attack surveillance into pattern classification problem and solves it with the ISVM. This method uses the ISVM to train and learn a huge mass of historical data in the industrial control network in order to secure the attack detection rule with which the real-time supervision will be enabled on the industrial control network. In the end, a replication experiment is cited with real data in power secondary network from a 110kv intelligent substation to reveal that the method features less supervision time-consuming and higher supervision precision.

1. Introduction

Chemical industry has high risks. Once the fire or explosion happens, property losses even casualties may be caused, such as the explosion accident broken out in Tianjin Binhai New Area in 2015. The power system works just like the lifeline of the whole production process of chemical companies, so that it is required to have high reliability for power supply to ensure that there are stable operation and normal production of chemical industry.

In recent years, information communication and automation control technologies have sprung up and find wide applications in every fields of production. In particular, the emergence of Industrial Control Networks (ICN) has made it possible to intelligently regulate and control the whole production process, seamlessly integrate the information and physical spaces. It has also enabled real-time centralized control on various field devices widely distributed in intelligent substations, power plants, etc. (Galloway et al., 2013), which not only improves the automation level of the power grid but also helps incessantly provide high-quality electric energy, thus catering to the safety expectations of chemical enterprises. Since its inception, the ICN more concerns how to ensure its availability, instantaneity and business continuity. Cybersecurity, however, has not yet been regarded as a core in design, thus causing frequent attacks against ICN in recent years. More frequently, grave consequences ensued (Amin et al., 2013; Vollmer et al., 2014; Won et al., 2008), for example, various viruses against ICN, such as Night Dragon, Duqu and Nitro, have been discovered, which attack against ICN concealed, resulting in serious consequences, since 2013 (Zou et al., 2013).

In response to the increasingly severe cybersecurity situation in the field of ICN, a cybersecurity attack detection system is proposed to improve the defense of ICN against cybersecurity attacks and ensure its safe and stable operation. This method abstracts attack supervision into a pattern classification problem and solves it using Improved Support Vector Machine (ISVM) which then trains and learns a huge mass of historical data in the ICN in order to obtain attack timely detection rule for the ICN. In the end, an experiment is conducted on field data from a 110kv intelligent substation power secondary network to test the method in terms of supervision time and precision (Cao et al., 2018).

2. ISVM

ISVM evolved from Support Vector Machine (SVM). Its core idea is still to map the target problem from Euclidean space to Hilbert space, making it a linearly separable dual problem, and then solving the optimal hyperplane, that is, the pattern classification for the target problem is done. The biggest difference between the two lies in that ISVM uses swarm intelligence optimization algorithms (such as genetic algorithm, ant

Please cite this article as: Shen H., Cheng F., Lu X., 2018, Information security of chemical substation based on isvm, Chemical Engineering Transactions, 71, 1135-1140 DOI:10.3303/CET1871190

colony algorithm, differential evolution algorithm, etc.) to dynamically search for optimal training parameters during the training process, thus make up the gap of SVM which works out a lower classification precision caused by random training parameters (Xu et al., 2012; Li et al., 2012; Dong et al., 2005). The working process of ISVM is shown in Fig. 1.

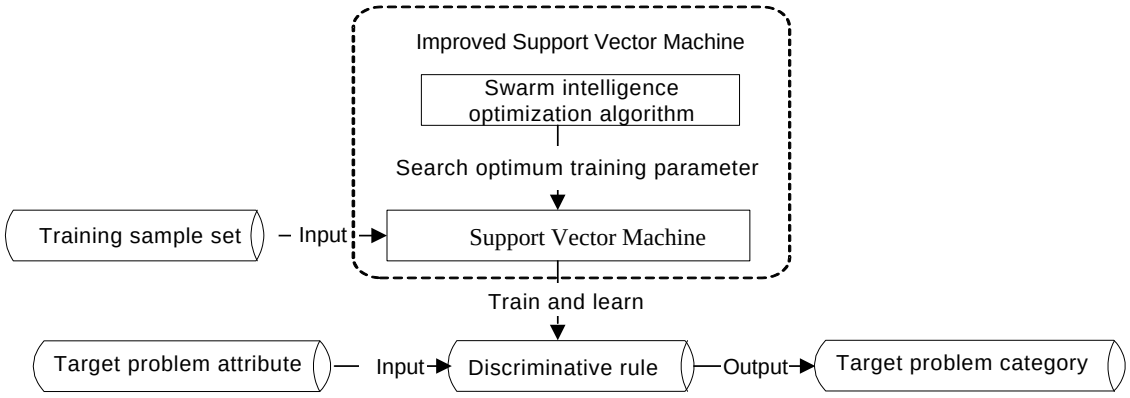

Figure 1: Working process of ISVM

3. Cybersecurity attack detector of industrial control network

3.1 Working principle

ICN is a typical "three-layer and two-network" structure consisting of station control layer, spacer layer, process layer, and station control layer network and process layer network, among which, the station control layer network uses TCP/IP communication, and the process layer network uses MAC frame for communication.

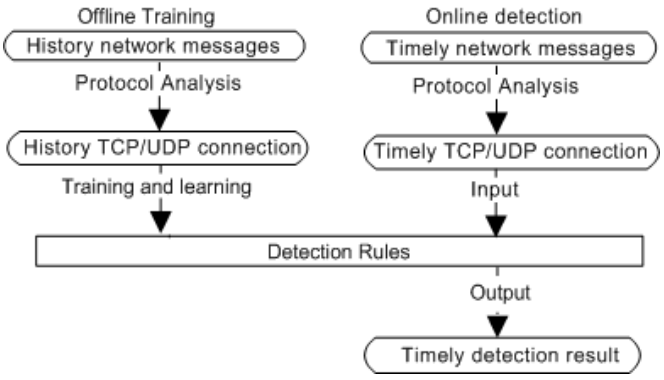

Figure 2: The work principle of cybersecurity attacks detection

Data sources of cybersecurity attack detector mainly concentrate on the messages in the station control layer network. Its working principle is such that the feature attribute of real-time TCP/UDP connection matches the detection rule after several network messages are restored into one TCP/UDP link, and eventually the attack detection result is available. The whole supervision process can be divided into two parts: offline training and online detection, as shown in Fig. 2. The historical TCP/UDP link in Fig. 2 contains both characteristic attributes and detector results, so that the training rules will be available by training and learning the historical TCP/UDP connections using the pattern classification method.

3.2 Work process of attack detector

Step 1, Data preparation

(1) Network message acquisition. Probes are deployed in the station control layer network of ICN to collect network messages.

(2) TCP/UDP connection restoration. The protocol analysis is used to restore the collected network messages into TCP/UDP connection to display the purposes of their original network behaviors. The TCP/UDP connection can be a normal operation command or a network attack.

(3) Data set generation. With the KDD Cup99 data set generation rule, the TCP/UDP connections constitute a data set (Hettich and Bay, 1999). Each sample with 42 attributes in the data set corresponds to a TCP/UDP connection. The first 41 attributes are used to represent the basic information of the TCP/UDP connection (such as basic and content features of TCP/UDP connections, time-based network traffic features, host-based network traffic feature, etc.), the attribute 42 represents the purpose of network behaviors (such as normal operation instructions or network attacks) that the TCP/UDP connections correspond to. The first 41 attributes are used as the input variables of the ISVM, while the feature attribute 42 is the output of the ISVM. Data set is divided into the training and test sample sets by the ratio of 1:1 as required by the open-set test (Hettich and Bay, 1999). The ISVM trains and learns these sample sets to get the initial detection rule and uses the test sample set to constantly correct the precision of the initial detection rule until the one that meets the desired threshold is available.

Step 2, Generation of detection rule

(1) Pattern classification algorithm selection and parameter setting. μ-SVM is chosen as the classification algorithm of ISVM. Particle Swarm Optimization (PSO) is used as the parameter optimization algorithm of ISVM. Kernel function type of μ-SVM adopts Gaussian kernel function, such as formula (1), x′ is the center point of Gaussian kernel function, and the penalty factor C and kernel width factor σ of Gaussian kernel function are dynamically optimized by PSO. Set the PSO population size to 5 and the maximum iterations to 300.

$$K(\mathrm{x},\mathrm{x}') = e^{\frac{-\|x-x'\|^2}{\sigma 2}} \tag{1}$$

(2) Training the set of samples. The μ-SVM trains and learns the set of samples to obtain the initial detection rule, and detect its Total Detection Accuracy (TDA) using the set of test samples, False Negative Rate (FNR) and False Alarm Rate (FAR). If the TDA is greater than or equal to 98%, the initial detection rule is just the last detection rule; otherwise, the training parameters of the μ-SVM are updated by PSO to generate new detection rules, and the TDA is then calculated until the TDA of a generation of detection rules is greater than or equal to 98%, or the PSO iteration reaches 300 generations, the sample set training is ended. Suppose the sample set has a normal network behavior samples (a_1 is correctly identified, a_2 is misidentified) and b cyberattack samples (b_1 are correctly identified, b_2 is misidentified), then the TDA, FNR and FAR are calculated by formulae (2)-(4). The generation process of the detection rules is shown in Fig. 3.

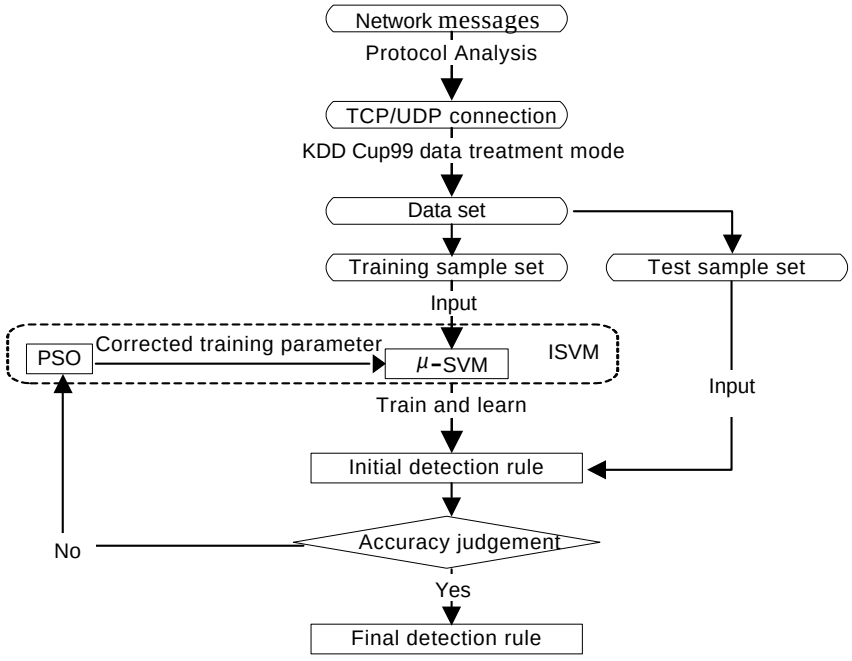

Figure 3: The generative process of detection rules

$$\mathrm{TDA}=\frac{a_1+b_1}{a+b}\times 100\% \tag{2}$$

$$FNR=\frac{b_2}{a+b}\times 100\% \quad (3)$$

$$FAR=\frac{a_2}{a+b}\times 100\% \quad (4)$$

Step 3, Online detection
The network messages are timely collected with the probes deployed in the ICN station control layer network, and restored to a real-time TCP/UDP connection according to the data pretreatment in step 1. At last, real-time TCP/UDP connection and detection rules match categories to get real-time detection results for information security attacks.

4. Experiment

4.1 Experimental environment

In order to test the availability of the ICN information security attack detector proposed herein, an experimental environment is set up in a 110kv intelligent substation, and an replication experiment is designed as follows: the attacker host and information security attach detector host are accessed to the station control layer network of the intelligent substation (built-in Wireshark and attack detection rule), in which the attacker host attacks against the station control layer network and its host, as shown in Table 1, and the information security attack detector host restores the network messages timely captured into the TCP/UDP network connections and matches them with the detection rules to give real-time alerts for information security attacks.

Table 1: The cybersecurity attacks in experiment

No.	Attack types	Attack target
1	DDoS	Station control layer network
2	IP scan	All hosts in station control layer
3	Port scan	All hosts in station control layer
4	Password detection	Operator training host
5	Buffer overflow	Intelligent substation detector host
6	Remote control	Operator training host

4.2 Experimental data

6,000 TPC/UDP connections are restored from the network messages captured by the information security attack detector host, including normal samples 900, distributed denial of service (DDoS) attack samples 2000, and scanning attack samples (IP and port scan) 2500, application attack samples 200 (password detection and buffer overflow), and false message attack samples 400 (remote control). The training and test sample sets are built according to the method in Section 3.2, as shown in Table 2.

Table 2: The experimental samples distribution

	Training sample set	Test sample set
Normal	466	434
distributed denial of service	989	1011
Scanning	1240	1260
Application	99	101
False messages	206	194

4.3 Experiment procedure and analysis of results

In order to fully demonstrate the performance advantages of the proposed method, the μ-SVM method with unoptimized training parameters is used for horizontal comparison of detection performance. Position change of PSO population particles in the training model based on ISVM is shown in Fig. 4. The accuracy variation of each generation of detection rules is shown in Fig. 5. The optimal training parameters of μ-SVM optimized by PSO are the penalty factor C=1430 and kernel width factor σ = 0.001, and the final detection results are shown in Table 3.

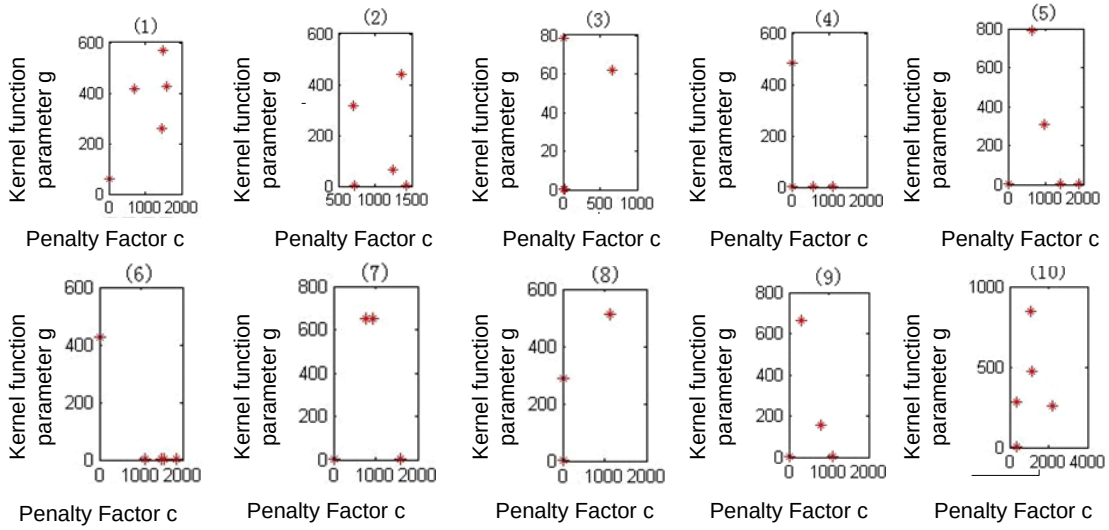

Figure 4: Changes in the particle positions for PSO

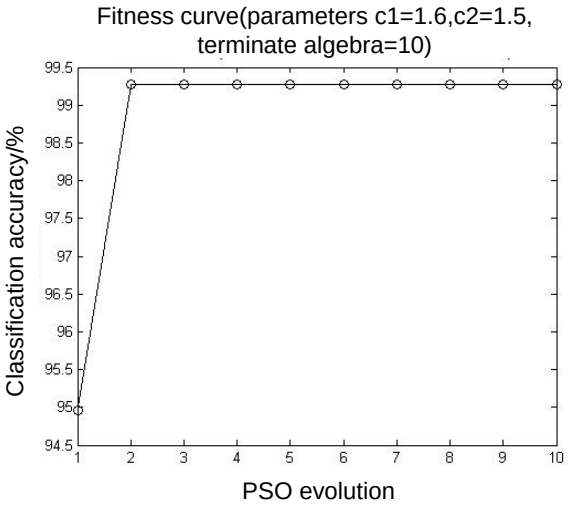

Figure 5: The accuracy changes of chronicles detection rules

Table 3: The confirming experimental results

		ISVM	*μ*-SVM
Normal	Correct recognition	434	434
	Wrong recognition	0	0
distributed denial of service	Correct recognition	1011	1004
	Wrong recognition	0	7
Scanning	Correct recognition	1256	1245
	Wrong recognition	4	15
Application	Correct recognition	96	96
	Wrong recognition	5	5
False messages	Correct recognition	181	118
	Wrong recognition	13	76
False negative rate		0.73%	3.43%
False alarm rate		0%	0%
TDA		99.27%	96.57 %
Detection time		33.92 s	33.92 s

Combining the above experiment diagram, the experimental results are analyzed:
(1) From the perspective of detection accuracy, the TDA of ISVM-based information security attack detector exceeds 99%, and its accuracy in every way is obviously superior to the horizontal comparison method since the ISVM uses PSO to dynamically search for μ-SVM training parameters, thus effectively improving the detection accuracy;
(2) From the perspective of detection time-consuming, the ISVM-based information security attack detector takes 33.92 s to detect 3,000 samples, that is, it detects one sample every 0.0113 s, which can meet the real-time detection requirements.
It follows that the ISVM-based information security attack detection model features higher detection accuracy and faster detection speed.

5. Conclusion

This paper proposes a cybersecurity attack detector applicable to the industrial control network to improve the active defense capacity against cybersecurity attack on industrial control networks. This method regards attack detection as a pattern classification problem and solves it by using ISVM to obtain detection rules with which real-time detection is enabled on industrial control network cybersecurity attacks. Simulation experiment on the 110kv intelligent substation environment reveals that this method can achieve high-precision surveillance against cybersecurity attacks on a short-time basis, and seems more reliable.
How to further improve the detection precision of this method, especially to reduce its False Negative Rate and improve its detection speed will be the focus of future study.

Acknowledgement

This paper is supported by Baoding science and technology research and development guidance project (14ZS005, 17ZN001).

References

Amin S., Litrico X., Sastry S.S., 2013, Cyber Security of Water SCADA Systems – Part II: Attack Detection Using Enhanced Hydrodynamic Models, IEEE Trans. On Control Systems Technology, 21(5),1679-1693, DOI: 10.1109/TCST.2012.2211874

Cao S., Li N., Yu N., 2018, Power electronic transformer control technology based on dsp, Chemical Engineering Transactions, 66, 1261-1266. DOI: 10.3303/CET1866211

Dong S.K., Nguyen H.N., Park J.S., 2005, Genetic algorithm to improve SVM based network intrusion detection system, Advanced Information Networking and Applications, AINA 19th International Conference on, 1(2), 155-158, DOI: 10.1109/AINA.2005.191

Galloway B., Hancke G.P., 2013, Introduction to Industrial Control Networks, IEEE Communications Surveys & Tutorials, 15(2), 860-880, DOI: 10.1109/SURV.2012.071812.00124

Hettich S., Bay S.D., 1999, KDD Cup 99 Task Description [EB/OL]., 1999, kdd.ics.uci.edu/databases/kddcup 99/task.html.

Li Y.C., Wang Y.F., 2012, A Misuse Intrusion Detection Model Based on Hybrid Classifier Algorithm, International Journal of Digital Content Technology and its Applications (JDCTA), 6(5), 25-33, DOI: 10.4156/jdcta.vol6.issue5.4

Vollmer T., Manic M., 2014, Cyber-Physical System Security with Deceptive Virtual Hosts for Industrial Control Networks, IEEE Trans. on Industrial Informatics, 10(2), 1337-1347, DOI: 10.1109/TII.2014.2304633

Won Y.J., Choi M.J., Hong W.K., Kim M.S., Hwang H., 2008, Fault Detection and Diagnosis in IP-Based Mission Critical Industrial Process Control Networks, IEEE Communications, 46(5), 172-180, DOI: 10.1109/MCOM.2008.4511666

Xu R.Z, Wang Y.F., 2012, Support vector regression based on Particle Swarm Optimization for distribution network theoretical line loss, Electric Power Automation Equipment, 32(5), 86-89, 93, DOI: 10.3969/j.issn.1006-6047.2012.05.016

Zou C.M., Zheng Z.Q., Liu Z.Y., Chen L.H., Chen M.C., 2013, Application of two power safety protection technology in industrial control system, Power grid technology, 37(11), 3227-3232.