International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 82

DETECTION OF FALSE DATA INJECTION ATTACK IN AUTOMATIC GENERATION CONTROL SYSTEM BASED ON LOCAL PREDICTOR AND SUPPORT VECTOR MACHINE

Z. Y. Chen, K. S. Xiahou, M. S. Li, Member, IEEE, T. Y. Ji, Member, IEEE, Q. H. Wu*, Fellow, IEEE

ABSTRACT

False data injection attack (FDIA) invades the automatic generation control (AGC) system and degrades the control performance, which may cause unstable operation of power system. Fast and accurate detection can help to reduce the impact of attacks. This paper presents a novel detection method, combining local predictor (LP) and support vector machine (SVM), for the FDIA of AGC system. The effects of different types of cyber attacks on AGC system are analyzed. The LP is applied to identify the local pattern of each point of historical data, and it extracts the information in a high dimension space with accurate predictions. The similar data obtained from LP are adopted to train the SVM, and the LP-SVM algorithm is presented to detect the attacks of AGC system. Simulation studies undertaken on a single-area AGC system reveal that the LP-SVM method outperforms traditional SVM and naive Bayes (NB).

Keywords: Automatic Generation Control System, False Data Injection Attacks, Local Predictor, Support Vector Machine

1. INTRODUCTION

Today's power grids are tightly connected through various communication systems, which makes the interconnect protection, monitoring and control intimate. One of the communication-related functions is the automatic generation control (AGC), which is performed by the supervisory control and data acquisition (SCADA) center [1]. The SCADA center of the power grid uses remote measurement devices and communication channels to collect useful data [2]. The power system controller of the SCADA center processes the collected data and sends the required commands to the relevant actuators. The AGC is a supplementary controller for the load frequency control (LFC) system, and it is the only automatic closed loop between the network and physical parts of the grid [3]. AGC improves the efficiency and quality of LFC operation. AGC is another target for economical scheduling of supplementary LFC loops. However, the dependence of AGC on communication makes the LFC system more vulnerable to the cyber attack.

False data injection attack (FDIA) is one of the most common types of cyber attacks on power grid, which combats AGC systems and may cause unstable of the power grid. The AGC system can locate FDIA and other types of intrusions through error data injection, such as denial of service, malware injection, spoofing and internal attacks [1]. In recent years, various studies have been conducted on the vulnerability of power systems to cyber attacks. For instance, in [4], the author designed a method that leverages the redundancy of measurement to mitigate the impact of an identified attack. Reference [5] discussed the effects of resonance attacks which is more complex FDIA, and it studied the performance of the LFC system. Reference [6] pointed out that destroying the stable power system through attacking the AGC system and proposed a program which is based on feedback linearization for detecting the optimal attack.

As to the issues, this paper presents a novel data-driven model for detecting attacks on AGC system, and it is based on local predictor (LP) and support vector machine (SVM). The SVM is a machine learning method based on statistical theory. The LP is applied to set up

Corresponding author.
E-mail address: wuqh@scut.edu.cn (Q. H. Wu).
Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).

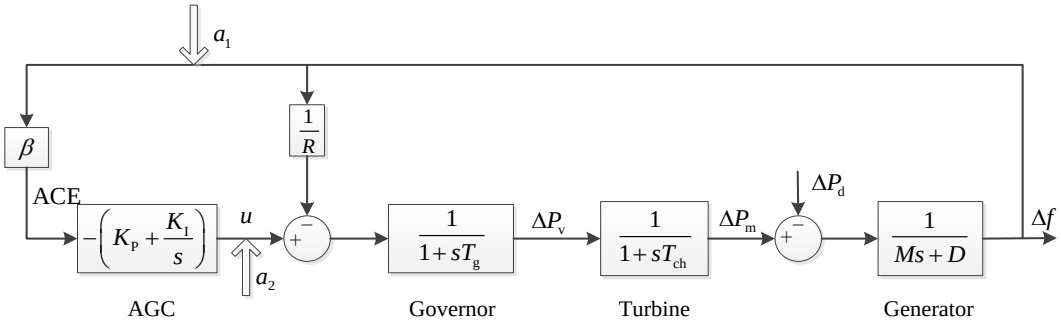


Fig. 1 Block diagram of a single-area LFC system.

training set from part of historical data and it identifies local pattern for each point. What is more, the LP can be displayed in a high-dimension space, it is obvious and intuitive.

Firstly, the dynamic model of a single-area LFC is built on the MATLAB/Simulink platform, and then various configurations of the model according to the various faults are set that may occur in an actual system. Thirdly, a series of fault signals are generated, which provide reasonable and sufficient data for the design of feature extraction and attacks detection methods. Lastly, the LP-SVM is tested in the dynamic model under different attacks and the results are compared with the SVM and naive Bayes (NB).

2. DYNAMIC MODEL OF LFC SYSTEM

Fig. 1 shows the dynamic model of a single-area LFC. It consists of three modules which are governor, turbine and generator respectively [7]. The generator correction is called area control errors (ACE) which is used for balancing the area. The proportional integral (PI) controller is the load frequency controller used in the current industry and is included in the model.

2.1 Structure of the LFC system

As shown Fig. 1, to model the dynamic characteristics, their transfer functions can be expressed as following [8], [9]. The relation between the frequency deviation Δf and power output ΔP_v is

$$\Delta P_v = -\frac{1}{R\left(1+sT_g\right)}\Delta f + u\,. \tag{1}$$

Where R is the speed regulation droop, ΔP_v valve position, and

$$T_g = \frac{1}{KR} \tag{2}$$

where K contains K_p and K_I, they represent proportional and integral gains respectively. The transfer function of a steam non-reheat turbines is

$$\frac{\Delta P_m}{\Delta P_v} = \frac{1}{1+sT_{ch}} \tag{3}$$

where ΔP_m is the deviation of the generator mechanical output and T_{ch} is the time constant of the turbine. The rotor and load can be used for the load frequency control.

Deduced as above, the dynamic model of one-area LFC system can be described as

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + F\Delta P_d \\ y(t) = C(x) \end{cases} \tag{4}$$

where

$$x(t) = [\Delta f, \Delta P_m, \Delta P_v, \int \mathrm{ACE}] \tag{5}$$

$$y(t) = [\mathrm{ACE}, \int \mathrm{ACE}]^T \tag{6}$$

$$A = \begin{bmatrix} -\frac{D}{M} & \frac{1}{M} & 0 & 0 \\ 0 & -\frac{1}{T_{ch}} & \frac{1}{T_{ch}} & 0 \\ -\frac{1}{RT_g} & 0 & -\frac{1}{T_g} & 0 \\ \beta & 0 & 0 & 0 \end{bmatrix} \tag{7}$$

$$B = [0, 0, T_g, 0]^T \tag{8}$$

$$F = \left[-\frac{1}{M}, 0, 0, 0\right]^T \tag{9}$$

$$C = \begin{bmatrix} \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

and ΔP_d, M, D, T_g and u denote the load, moment of inertia of the generator, generator damping coefficient, time constant of the governor, time constant and control signal of the turbine, respectively. In single-area LFC system, the ACE is defined as

$$\mathrm{ACE} = \beta\Delta f \tag{11}$$

where β is frequency bias factor.

2.2 Type of Attack

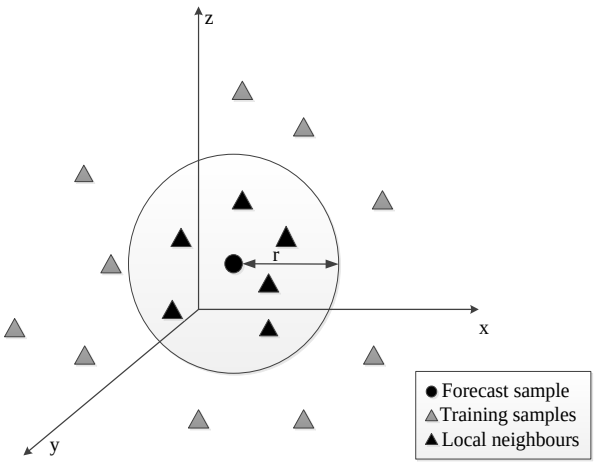


Fig. 2 The local neighbours are used in LP.

As shown in the Fig. 1, a_1 is the transmission channel of attack measurement data and a_2 is the transmission channel of attack control command. In the smart grid, false data injection is one of the most widespread types of attacks. Attackers inject different erroneous data on different signals on the communication system and lead to erroneous decisions, which results in huge damage to power system. The former model of the attack template discussed is presented in the previous subsection. In this subsection, we define several types of attack, t is time, t_{a} represents the attack period and Z_{mea} is the measurement and Z_{rea} is the real value.

2.2.1 Pulse attack

During the entire attack period, the measurements are set to higher/lower values. This type of attack modifies the measurement by a short pulse of time interval with attack parameter a_{p}.

2.2.2 Ramp attack

Ramp attacks modify measurement by adding $a_{\mathrm{r}} \cdot t$, a ramp function is increase/decrease over time gradually, where a_{r} represents the factor of ramp attacks. We can define a system of equations

$$Z_{\mathrm{mea}} = \begin{cases} Z_{\mathrm{rea}} & \forall \quad t \notin t_{\mathrm{a}} \\ Z_{\mathrm{rea}} + a_{\mathrm{r}} \cdot t & \forall \quad t \in t_{\mathrm{a}}. \end{cases} \tag{12}$$

2.2.3 Scaling attack

The scaling attack modifies the measurement higher/lower values by scaling attack parameters, which is a_{s}. We define a system of equations

$$Z_{\mathrm{mea}} = \begin{cases} Z_{\mathrm{rea}} & \forall \quad t \notin t_{\mathrm{a}} \\ (1 + a_{\mathrm{s}}) * Z_{\mathrm{rea}} & \forall \quad t \in t_{\mathrm{a}}. \end{cases} \tag{13}$$

3. LOCAL PREDICTOR AND SUPPORT VECTOR MACHINE

In this paper, the fault diagnosis method based on LP and SVM is proposed. The LP is applied to set up training set from part of historical data and it identifies local pattern for each point. What is more, the LP can be displayed in a high-dimension space, it is obvious and intuitive.

3.1 Structure of the LFC system

LP is different from global prediction, which selects partial historical data to set up training set. LP distinguish each point in the local pattern. This is the only one point and can be demonstrated in a more obvious way and a high dimensional space. Obviously, LP has a more purposeful modeling level and results in higher precision than global prediction.

3.1.1 Principle of Local Prediction

Take the three-dimensional phase space as an example. The mechanism of the local predictor is shown in Fig. 2. The original time series is placed in a three-dimensional phase space to form a batch of samples. LP selects a group of nearest neighbours, which are defined as local neighbours. Local neighbours have smaller Euclidean distance from the forecast sample and have a high similarity to the forecast samples.

3.1.2 The Structure of Local Predictoin

Collect the historical load data in time sequences. Single variable time series data contains all the information relative to variable. We can also understand that one-dimensional time series can be viewed as a lower dimensional compression with high dimensional information. In a high-variance system, using the embedding theorem to reconstruct a new space called phase space, the information can be extracted from a time series from on dimension to a higher dimension [10]. In the time domain, the sample $l(t)$, $t = 1, 2, \cdots, S$ indicates local information and the S is the length of time series. Accordingly, we can reconstruct the phase space through the delay coordinates and the function is

$$\mathbf{l}_t = [l(t), l(t+\tau), \cdots, l\left(t + (d-1)\tau\right)]^{\mathrm{T}} \tag{14}$$

where d represents the embedded dimension, τ is the time delay constant. Each $\mathbf{l}_t$ is unique in phase space and all $\mathbf{l}_t$ form a matrix together, which is written as

$$\mathbf{L}=[\mathbf{l}_1,\mathbf{l}_2,\cdots,\mathbf{l}_q]^{\mathrm{T}} \tag{15}$$

where $\mathbf{L}$ is a set of nearest neighbours has high similarity to the forecast sample in the phase space and $q=S-(d-1)\tau$.

3.2 Support Vector Machine Method

Support vector machine is a new machine learning method based on statistical theory, which maps the input sample space to the high-dimensional linear feature space through a nonlinear kernel function, and it is able to handle nonlinear regression problems. It overcomes the shortcomings of artificial neural network, which shows long training time, poor generalization ability, falling into local minimum easily and so on. Support vector machine improves the generalization ability of the model.

A brief description of the SVM algorithm is provided next. Get a training set $G=\{x_i,y_i\}_{i=1}^{S}$ with input vectors $x_i=\left(x_i^{(1)},\cdots x_i^{(S)}\right)^{\mathrm{T}}\in B^S$ and target labels $y_i\in\{-1,+1\}$, the support vector machine classifier, on the basis of Vapnik's [11] primitive description, it has to meet the following requirements

$$y^*(t)=\begin{cases}\omega^{\mathrm{T}}\phi(x_i)+b\geq=+1, & \text{if} \quad y_i=+1\\ \omega^{\mathrm{T}}\phi(x_i)+b\leq=-1, & \text{if} \quad y_i=-1.\end{cases} \tag{16}$$

In another way

$$y_i\left[\omega^{\mathrm{T}}\phi(x_i+b)\right]\geq 1, \qquad i=1,\cdots,S \tag{17}$$

where ω is the weight vector and the b represents the bias. Nonlinear function $\phi(\cdot):B^S\rightarrow B^{S_k}$ embeds the input or measurement space into a high-dimensional.

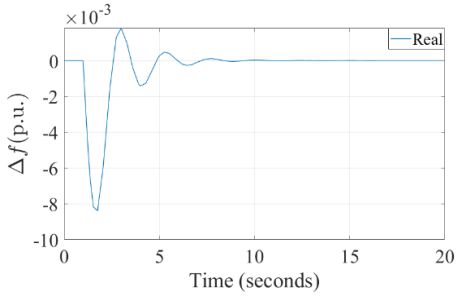
4. SIMULATION STUDIES

The formal model of the attack template is presented in the previous subsection. In this section, the impact of different types of attacks on the simulation model are observed and the accuracy of the novel method for detecting the attacks are verified.

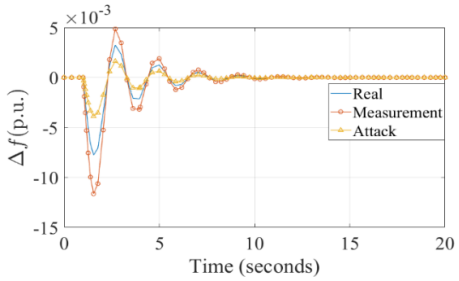
4.1 Impacts of FDIAS on System

4.1.1 Frequency deviation

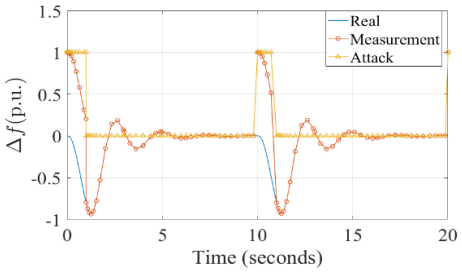
The following observations can be drawn from Fig. 3. Under normal circumstances, the real frequency value is disturbed by the load with a step in the first second, and after a period of fluctuation, the stability is achieved. After adding the scaling attack with a factor of 0.5, as shown Fig. 3(a) and Fig. 3(b), the real frequency value and the the frequency measurement value are biased and oscillation duration becomes longer. Comparing Fig. 3(c) and Fig. 3(a), as the pulse attack changes periodically, the real frequency value and the frequency measurement fluctuate within a periodic range and tend to stabilize.



(a) Under normal circumstances



(b) A scaling attack



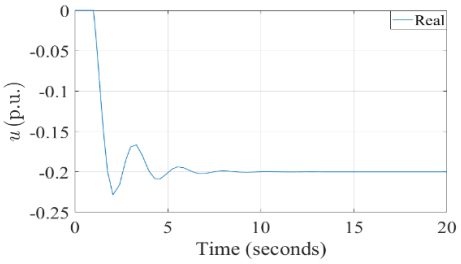
(c) A pulse attack

Fig. 3 Frequency changes under different attacks.

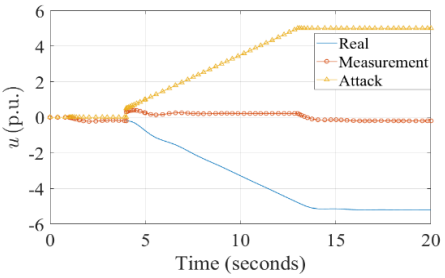
In modern power system, large frequency fluctuations can cause irreversible damage to equipment. But, the goal for this type of attackers is to initiate the frequency drop rapidly and trigger the load shedding scheme. In this case, attackers usually chose the scaling attack to change the frequency as soon as possible.

4.1.2 Control input

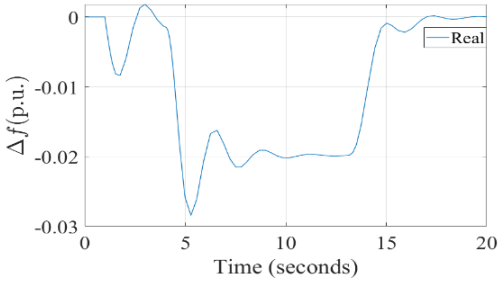
Under normal circumstances, the control signal is disturbed by the load, and after a period of fluctuation, the stability is achieved. As shown in Fig. 4(a) and Fig. 4(b), after adding the ramp attack with a slope of 0.5, as the attack signal increases, the real signal weakens rapidly and the measured signal tends to stabilize after a slight fluctuation. As to frequency signal, we can observe the apparent fluctuation and substantial decrease, then revert to original value after 18 seconds.



(a) Under normal circumstances



(b) A ramp attack



(c) Frequency deviation

Fig. 4 The influences of a ramping attack on control input.

The choices of attack parameters and attack methods are critical to the attackers. The right choices not only produce the desired impact, but also do not touch off the alerts in the control center.

4.1.3 Resonance attacks

In this scenario, firstly, the attackers steal the output information of power grid. Then, the attackers modify or fake the input based on resonance source which are chosen ahead of the time and send the unreal signals input to the target power grid [5]. Thirdly, if the input is within reasonable range, the input will be accepted by the target power grid and the attackers can get their goals include adjusting the generator. Comparing Fig. 3(a) and Fig. 5, after adding a pulse attack into the system in the fourth second, the load ΔP_d and frequency deviation fluctuate with the period of the pulse attack, but the frequency attack measurement is not changed at all.

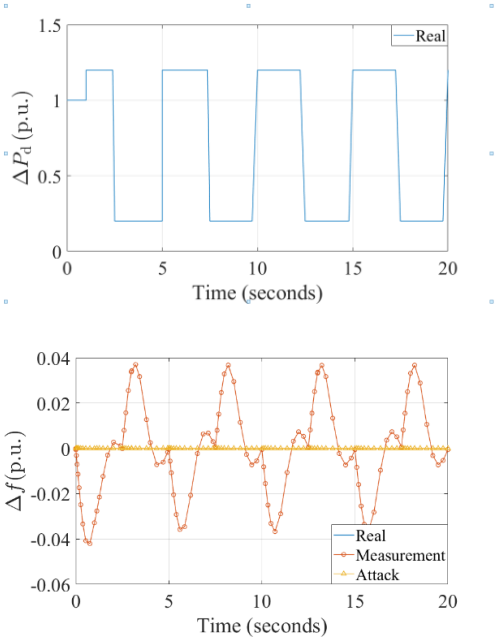


Fig. 5 Resonance attack on AGC.

4.2 Performance Evalution

In order to verify the efficiency of the proposed method, the detect results will be revealed and discussed below. First of all, the dynamic system of a single-area LFC was established in MATLAB/Simulink. Then, 7 types of faults are set in 3 locations. Among these attacks, scaling attack, ramping attack and pulse attack are set in the modules which are frequency measurement attack and control attack, and simulate pulse attack at the port of load input. Specifically, the type and target of each fault are shown in Table I. Then, 1000 samples are saved at each different fault, which are 7000 samples in total. Thirdly, 800 training samples and 200 testing samples are selected at every scenarios respectively. After the training and detecting processes are finished, the accuracy of three method for testing the types of faults are presented in Table II.

TABLE I: Type and target of each attack

Type	Target	Attack
I	Measurement channel	Scaling
II	Measurement channel	Ramp
III	Measurement channel	Pulse
IV	Control channel	Scaling
V	Control channel	Ramp
VI	Control channel	Pulse
VII	Load input	Pulse

TABLE II: Performance evaluation of the identification Method

Method / Accuracy / Fault	LP-SVM	SVM	NB
I	1.000	0.780	0.730
II	1.000	1.000	1.000
III	0.980	0.980	1.000
IV	1.000	1.000	0.980
V	1.000	1.000	1.000
VI	1.000	0.420	0.890
VII	1.000	0.000	1.000
Average	0.997	0.814	0.819

From Table I, comparing the detection method which is based on LP-SVM with SVM and NB demonstrates grate performance for all faults in all fault time series. To be more precise, the accuracy of LP-SVM has an average accuracy of 0.997 in seven different types of attacks. At the same time, the probability of fault I, II, IV, V, VI, VII are detected correctly is close to 1, which shows the LP-SVM has strong robustness. These advantages attributes to the fact that the LP can be exhibited in high dimensional space, and the training set is extracted based on partial historical data and each point I distinguished in the local pattern.

5. CONLUSION

In this paper, we have investigated the influences of FDIA on AGC system. Through various typical attack instances, the scaling, ramp and pulse attacks affect the stability of the power system, and attackers can destroy the electricity subtly under specific attacks. We have proposed one novel data-driven method for detecting attacks on a dynamic AGC system, based on LP and SVM. Results obtained from simulations demonstrate the higher accuracy and more stable performance for detecting attacks based on LP-SVM than SVM in various situations. Comparing with NB, the advantage of LP-SVM is also quite obvious. Moreover, under seven different attacks, the accuracy for detecting the attack can be maintained to 1, which indicates strong robustness.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China under Grand U1866210 and China Postdoctoral Science Foundation under Grand 2018M643082.

REFERENCE

[1] Manimaran Govindarasu, Adam Hann, and Peter Sauer. Cyber-physical systems security for smart grid. Future Grid Initiative White Paper, PSERC, Feb; 2012.

[2] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven Weller, and Z. Y. Dong. A review of false data injection attacks against modern power systems. IEEE Transactions on Smart Grid, 2016: 1–1.

[3] Yee Wei Law, Tansu Alpcan, and Marimuthu Palaniswami. Security games for risk minimization in automatic generation control. IEEE Transactions on Power Systems 2015; 30(1):223–232.

[4] Rui Tan, HoangHai Nguyen, Eddy. Y.S. Foo, David K.Y. Yau, Zbigniew Kalbarczyk, Ravishankar K. Iyer, and Hoay Beng Gooi. Modeling and mitigating impact of false data injection attacks on automatic generation control. IEEE Transactions on Information Forensics & Security 2017, 12(7): 1609-1624.

[5] Yong dong Wu, Wei Zhuo, Weng Jian, Li Xin, and Robert H. Deng. Resonance attacks on load frequency control of smart grids. IEEE Transactions on Smart Grid, 2017:1–1.

[6] Peyman Mohajerin Esfahani, Maria Vrakopoulou, Kostas Margellos, John Lygeros, and Göran Andersson. A robust policy for automatic generation control cyber attack in two area power network. In 49th IEEE Conference on Decision and Control (CDC). IEEE 2010; 5973–5978.

[7] L. Jiang, W. Yao, J. Y. Wen, S. J. Cheng, and Q. H. Wu. Delay-dependent stability for load frequency control with constant and time-varying delays. IEEE Transactions on Power Systems. 2012;27(2):932–941.

[8] Prabha Kundur, Neal J Balu, and Mark G Lauby. Power system stability and control, volume 7. McGraw-hill New York; 1994.

[9] Hassan Bevrani. Robust power system frequency control. Springer; 2009.

[10] Pablo L. De NSpoli, Irene Drelichman, and Nicolas Saintier. Weighted embedding theorems for radial besov and triebel-lizorkin spaces. Mathematics 2016;233(1).

[11] Vladimir N Vapnik. Statistical learning theory. Annals of the Institute of Statistical Mathematics. 2003; 55(2):371–389.