Detection of Power Data Tampering Attack based on Gradient Boosting Decision Tree

Guo Ping1,*

1 Liaoning Jianzhu Vocational College, Liao Yang City, Liao Ning Province, 111000, China

*Corresponding author's e-mail: guoping3360@126.com

Abstract. In order to improve the detection accuracy of the data tampering attack of the Power Cyber-physical System, a method of Power CPS attack detection based on Gradient Boosting Decision Tree is proposed in this paper from the angle of artificial intelligence. Firstly, by setting up the iForest anomaly value equation of physical system and IDS Shannon entropy function of information system, the feature extraction of data tamper detection is realized. Then, the CART decision tree is taken as the base learner of the detection classification, the target is minimized by the loss function, and the high intensity attack detection model is designed by iterative combination. Finally, the three-dimensional adaptive chaotic FOA algorithm is proposed for dynamic optimization of model parameters, an attack detection model under training optimal parameters. The analysis of the example shows that show that the proposed method can effectively detect power CPS data tampering attacks and has excellent detection precision.

1. Introduction

With "Ukraine blackout" incident, the information security problem of power system has risen to a new level[1]. The new type of power CPS network attack is called network cooperative attack[2]. Attackers use carefully designed and diversified attack methods, make attacks begin in the information domain and specifically target in the physical domain, so as to maximize the loss of power system[3-4]. Data tampering attack, as a kind of network cooperative attack, may make the key equipment in the power system stop working or perform wrong operation, leading to large-scale power failure, casualties and other major power accidents. For the attack detection of power CPS, relevant scholars combine the access control, authentication encryption, intrusion detection and other technical means in the field of information security with the original power grid security technology, and have made great progress in practicability and accuracy[5]. However, with the further development of power system, it is difficult for the current detection methods to obtain high precision and high efficiency decision results when dealing with large-scale power data and covert data tampering attacks[6-8].

In recent years, breakthroughs have been made in the research of big data, data mining and artificial intelligence, providing new ideas for solving such problems. From the perspective of the new generation of artificial intelligence algorithms, this paper proposes a power CPS data tampering attack detection method based on gradient lifting decision tree. On the one hand, the iForest (Isolation Forest) anomaly score equation of physical system and the Shannon entropy function of information system IDS are established to realize the feature extraction of data tampering attacks. On the other hand, a high-intensity classification model of attack detection was iteratively constructed by combining
decision tree with gradient lifting, and a three-dimensional adaptive chaotic fruit fly algorithm was
designed to dynamically optimize the hyperparameters of the model.

2. Feature extraction of power CPS data tampering attack
The data tampering attack of power CPS is to illegally tamper the measurement data of physical nodes
by invading the information system, so as to malfunction the key equipment and mislead the operation
maliciously. It will inevitably leave traces on the data level of the information system and physical
system. Therefore, according to the abnormal characteristics of the measured data in the physical
system and the alarm characteristics of the IDS data in the information system, the data tampering
feature extraction is realized comprehensively.

2.1. Abnormal score feature extraction of physical system iForest
Power SCADA data is taken as the subject of physical data, and the goal of feature extraction is to
quantify data tampering behavior into abnormal features. Due to the traditional state estimation data
identification method exists residual pollution easily and residual annihilation phenomenon, this article
is based on isolated forests algorithm, and establish a physical data iForest anomaly score equations, to
realize the feature extraction of a physical system.

The establishment of iForest is made up of multiple isolated trees iTrees, which is a random binary
tree. The establishment process is as follows:
1) Feature $A$ was randomly selected from physical data set $D_p$;
2) Random selection of a single value of the feature $V$;
3) According to the characteristics of the selected $A$ binary tree split on each record, if any of the
property $A$ record $R < V$, put the record in the left child node, if $R \geq V$, on the right child nodes;
4) The left child node and the right child node are recursively constructed until each sample is
isolated or the height of the tree $L$ reaches the limit height to form iTree. By quantifying the
traversal depth of tested sample $X$ in each iTre e, the following anomaly score quantization
equation is defined:

$$ H(t) = \ln(t) + \xi $$

(1)

The iForest anomaly score of each physical data $x$ can be expressed as:

$$ iscore(x) = \frac{-E[h(x)]}{\mu} $$

(2)

Where, $\xi$ is Euler's constant, $h(x)$ is the path length of $x$, i.e. the sum of the edges from the root node to
the isolated node, and $E[h(x)]$ is the mean of the path length on all iTrees. When $iscore(x)$ approaches
0.5, the higher the normality, and when it approaches 1, the higher the abnormality.

Then the physical characteristic of attack detection of the i-th power device is:

$$ p_i = [ID_i, iscore(x, i), f_1, f_2, ..., f_n] $$

(3)

2.2. Data tampering attacks feature data sets
In modern power system, the network communication equipment on each measured line has a unique
IP address in the network space, and the number of information nodes is much greater than that of
physical nodes, and there are corresponding ways of interdependence between the physical network
and the information network.

If the i-th physical node is monitored by the information node of S station, the information system
and the physical system can be mapped by connecting the IP address of the network device with the
device ID of the physical node. The data tampering attack characteristic of the i-th power device is.
\[(P, C) = (p_i, H_{ave}(Q_s), H_{ave}(Q_s), H_{ave}(L))\]  \hspace{1cm} (5)

Where, \((P, C)\) is the attack detection feature at physical node \(i\), \(H_{ave}(Q_s), H_{ave}(Q_s), H_{ave}(L)\) are the mean values of source IP threat entropy, destination IP threat entropy and datagram length Shannon entropy of network equipment of \(S\) station respectively. \((P, C)\) represents the attack characteristic data set of all systems.

The feature data set is normalized to integrate network layer and physical layer features in the same data space and eliminate the influence of outlier data on attack detection. The normalization formula is shown as follows:

\[y_{i,j} = \frac{x_{i,j} - \min(x, \gamma)}{\max(x, \gamma) - \min(x, \gamma)}\] \hspace{1cm} (6)

3. Attack detection model based on gradient lifting decision tree

The advantages of this model lie in higher classification accuracy compared with other single classification models, and better generalization ability and construction efficiency compared with artificial neural network.

Now the attack characteristic data set containing \(N\) training samples is given:

\[D = \{(X_1, Y_1), (X_2, Y_2), ..., (X_N, Y_N)\}\]

\[X = \{C, P\}\]

\[Y \in \{0, 1\}\] \hspace{1cm} (7)

Where, \(X = (X^{(1)}, X^{(2)}, ..., X^{(N)})^T\) is the attack characteristic data set, \(Y_i \in \{0, 1\}\) is the class label, that is, whether it is attacked or not. The learning objective of the model is to find the optimal gradient lifting decision tree model \(F_{boost}(x)\) according to the given data tampering characteristic data set, so that the loss function \(L(y, F(x))\) mapping from \(x\) to \(y\) can be minimized, so that the model can correctly judge whether the power data samples are subjected to data tampering attack.

Each upgrade of the model is to reduce the residual in the direction of the minimum value of the loss function of the previous generation model, and the classifier with higher accuracy is constantly established until the number of iterations is satisfied.

1) Define the residual \(r_{im}\) as the value of the negative gradient of the loss function in the current model:

\[r_{im} = - \left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F_m(x) = F_{m-1}(x)} \] \hspace{1cm} (8)

2) If the number of iterations is \(M\), the estimated residuals obtained in the above equation are taken as input to obtain the leaf node region \(R_{nm}\) of \(M\) decision trees, where \(n = 1, 2, ..., N\).

\[R_{nm} = \arg \min \sum_{i=1}^{N} [r_{im} - F(x_i)]^2\] \hspace{1cm} (9)

3) Formula 10 can give the optimal step \(\beta_{nm}\) about loss function gradient descent direction, make minimum loss function:

\[\beta_{nm} = \arg \min_{\beta} \sum_{x \in R_{nm}} L(y_i, F_{m-1}(x) + \beta)\] \hspace{1cm} (10)

4) Build higher precision of weak classifier model \(F_{boost}(x)\), define \(\nu \in (0, 1]\) is learning rate, in the model to avoid over fitting:

\[F_m(x) = F_{m-1}(x) + \sum_{n=1}^{N} \nu \beta_{nm} I (x \in R_{nm})\] \hspace{1cm} (11)

5) At the end of the iteration, the final gradient lifting decision tree model was obtained by combining \(M\) weak classifiers with higher precision:
\[ F_{\text{boost}}(x) = \sum_{m=1}^{M} \sum_{n=1}^{N} \beta_{nm} I \quad (x \in R_{nn} ) \quad (12) \]

### 4. 3D adaptive chaotic fly optimization algorithm

The processing steps of 3D Adaptive Chaotic Drosophila Algorithm (V3ACFOA) is as follows:

1) Initialization the number of iterations Maxgen, flies population size Sizepop, and the initial position \( X_{axis}, \ Y_{axis} \) and \( Z_{axis} \), fitness variance threshold \( \mu \), chaotic number \( T \).

2) The three-dimensional random direction and distance of Drosophila individuals were defined, and the smell was used to search for food. \( RandomValue \) was the random distance.

\[
\begin{align*}
X_i &= X_{axis} + RandomValue \\
Y_i &= Y_{axis} + RandomValue \\
Z_i &= Z_{axis} + RandomValue \\
\end{align*}
\quad (13)
\]

3) By flies in the distance calculation formula to fly the distance from the origin \( Dist_i \), distance from bottom of \( Si \), as the taste of fruit flies concentrations determined value, and use the fitness \( Function \) and taste concentration determination value \( Si \), please get the flavor of the fruit flies location concentration \( Smelli \): to solve the population in all taste concentration \( Smelli \) odor concentration in the flavor of the largest individual.

4) Obtaining the maximum \( bestSmell \) and its three-dimensional \( X, Y \) and \( Z \) coordinates, so that the Drosophila population flies to this position.

\[
\begin{align*}
\text{Smellbest} &= bestSmell \\
X_{axis} &= x(\text{bestindex}) \\
Y_{axis} &= y(\text{bestindex}) \\
Z_{axis} &= z(\text{bestindex}) \\
\end{align*}
\quad (14)
\]

5) By type (15) to calculate the average concentration \( Smell_{avg} \) taste, according to the type (16) to calculate the concentration variance \( \sigma^2 \).

\[
\begin{align*}
\text{Smell}_{avg} &= \frac{\sum_{i=1}^{\text{SizePop}} \text{Smell}_i}{\text{SizePop}} \\
\sigma^2 &= \frac{\sum_{i=1}^{\text{SizePop}} (\text{Smell}_i - \text{Smell}_{avg})^2}{\text{SizePop}} \\
\end{align*}
\quad (15, 16)
\]

6) Algorithm of adaptive process, if \( \sigma^2 < \mu \) and \( T > 0 \), introducing chaos search, the type of chaos search a way of evolution:

\[
C_{x_{t+1}} = 4C_x (1 - C_x) \quad (17)
\]

Where, \( C_x \in [0,1] \), the optimization parameters \( x \in [a_i, b_i] \), the mapping can be done back and forth through the following equation

\[
\begin{align*}
C_x &= (x_i - a_i) / (b_i - a_i) \\
x_i' &= a_i + C_x (b_i - a_i) \\
\end{align*}
\quad (18, 19)
\]

In the chaotic optimization process of the algorithm, the optimal individual position variables \( X_i, Y_i \) and \( Z_i \) of Drosophila were transformed into chaotic variables \( CX_i, CY_i \) and \( CZ \) by equation (18), and the chaotic variables were evolved by equation (17). Finally, the chaotic variables were transformed into a new position \( X_i', Y_i' \) and \( Z_i' \) in the optimized region by equation (19). Set \( T = T - 1 \) and go to step 3).

otherwise, \( \sigma \geq \mu \) or \( T = 0 \) go to step 7).
7) Iterative optimization is performed to solve the optimal parameters, and steps 2) ~ 6) are repeated. When the optimization results meet the accuracy requirements or the termination condition \( g = \text{Maxgen} \), the algorithm stops.

### 5. Calculation examples and analysis

In order to verify the accuracy and efficiency of the feature extraction method, two other non-linear and linear feature extraction methods, i.e. local linear embedding method (LLE) and principal component analysis method (PCA), are selected to compare with the feature extraction method of iForest-LLE proposed in this paper. The training set and data set are reduced to the specified dimension respectively, and the 304 dimension measurement data samples generated in IEEE118 bus node system by gbdt are used for model training and attack detection.

![Figure 1. Model training time of different feature extraction methods](image1)

One of the purposes of feature extraction is to reduce the dimension of data, if the dimension is too large, it will not help the training of attack detection model. The specified dimensions discussed in this article are below 30. It can be seen in Figure 1. With the increase of dimensions, the training time of PCA increases the fastest and basically linearly. In this paper, the model training time of the iForest-LLE feature extraction method is faster than that of the basic LLE method, and the training time of the two methods is more stable with the increase of the specified dimension.

![Figure 2. Attack detection accuracy of different feature extraction methods](image2)

Figure 2 shows the attack detection accuracy (obfuscation matrix accuracy) of different dimensionality reduction methods. Compared with the basic LLE and PCA feature extraction methods, the feature extraction method of iForest-LLE proposed in this paper has more advantages in detection accuracy, When the specified dimension is 6 dimensions, the attack detection accuracy reaches 0.9523,
when the specified dimension is 10 dimensions, the attack detection accuracy reaches the highest 0.953, when the LLE method is 10 dimensions, the detection accuracy reaches the highest 0.925, while when the PCA method is 15 dimensions, the detection accuracy reaches the highest 0.905.

In conclusion, the iForest-LLE feature extraction method proposed in this paper has a great advantage in the training time and detection accuracy of attack detection model. Considering the training time and detection accuracy in a balanced way, the best training effect can be achieved when the specified dimension is 6 dimensions.

6. conclusion
In this paper, an attack detection method based on gradient lifting decision tree is proposed for power CPS data tampering attack, and a three-dimensional adaptive chaotic fruit fly algorithm is designed to solve the parameters. By simulating the data tampering attack, the experiment verifies that the proposed method can effectively detect the attack in the form of the probability of being attacked. Meanwhile, compared with the other three methods, the proposed method has higher accuracy rate, higher precision rate and lower missed and false detection rate, which achieves the target of high precision detection.

The research in this paper provides a new idea for power CPS network attack detection method from the perspective of artificial intelligence, but focuses on SCADA data tampering and malicious control attack with high degree of damage on transmission side, and the form of feature extraction elements and detection results is relatively single. Future research can start from many kinds of collaborative attack scenarios, such as WARMS system and AMI system, it can provide more comprehensive attack detection results.

Acknowledgment
This work is supported by Science Research Project of Education Department of Liaoning Province (JYT-L202005).

References
[1] ZHAO Junhua, LIANG Gaoqi, WEN Fushuan. et al. Lessons Learnt from Ukrainian Blackout: Protecting Power Grids Against False Data Injection Attacks[J]. Automation of Electric Power Systems, 2016, 40(7):149-151.
[2] GUO Qinglai, XIN Shujun, SUN Hongbin. et al. Power System Cyber-physical Modelling and Security Assessment: Motivation and Ideas[J]. Proceedings of the CSEE, 2016, 36(6):1481-1489.
[3] LIU Nian, YU Xinghuo, ZHANG Jianhua. Coordinated Cyber-attack: Inference and Thinking of Incident on Ukrainian Power Grid[J]. Automation of Electric Power Systems, 2016,40(06):144-147.
[4] Jia, K.; Wang, Z.; Fan, S. Data - centric approach: a novel systematic approach for cyber physical system heterogeneity in smart grid. IEEE Transactions on Electrical and Electronic Engineering. 2019, 14(5), 748-759.
[5] Lai, K.; Illindala, M.; Subramaniam, K. A tri-level optimization model to mitigate coordinated attacks on electric power systems in a cyber-physical environment[J]. Applied energy, 2019, 235, 204-218.
[6] Haes, Alhelou, H.; Hamedani-Golshan, M. E.; Njeta, T. C. A survey on power system blackout and cascading events: Research motivations and challenges. Energies, 2019, 12(4), 682.
[7] Dai, Q.; Shi, L.; Ni, Y. Risk Assessment for Cyberattack in Active Distribution Systems Considering the Role of Feeder Automation. IEEE Transactions on Power Systems. 2019, 34(4), 3230-3240.
[8] Lee, L.; Hu, P. Vulnerability analysis of cascading dynamics in smart grids under load redistribution attacks. International Journal of Electrical Power & Energy Systems. 2019, 111, 182-190.