

# Big Data and Knowledge Graph Based Fault Diagnosis for Electric Power Systems

Yuzhong Zhou<sup>1</sup>, Zhengping Lin<sup>1</sup>, Liang Tu<sup>1</sup>, Yufei Song<sup>2</sup>, and Zhengrong Wu<sup>2</sup>

<sup>1</sup>Electric Power Research Institute of China Southern Power Grid, Guangzhou, Guangdong, China.  
(e-mail: yuzhong\_zhou@hotmail.com, zhengping\_lin@hotmail.com, Liang\_Tu1@hotmail.com).

<sup>2</sup>China Southern Power Grid, Guangzhou, Guangdong, China.  
(e-mail: zhengrongwu@126.com, yufei\_song1@126.com).

## Abstract

Fault detection plays an important role in the daily maintenance of power electric system. Big data and knowledge graph (KG) have been proposed by researchers to solve many problems in industrial Internet of Things, which also give lots of potentials in improving the performance of fault detection for electric power systems. In particular, this paper analyzes a distributed knowledge graph framework for fault detection in the electric power systems, where multiple devices train their local detection models used for fault detection assisted with a central server. Each device owns its local data set composed of historical fault information and current device state, which can be used to train a local model for fault detection. To enhance the detection performance, the distributed devices interact with each other in the KG framework, where the devices ought to achieve the regional computation in addition to the model aggregation within a specified latency threshold. Through searching for the vibrant qualities together with determined ability at the devices, we enhance the knowledge graph framework by the optimum variety of energetic devices together with the restriction of latency as well as data transmission. Particularly, two data transmission frequency allocation (FA) schemes are developed for the distributed knowledge graph framework, through which *scheme I* is actually based after the instantaneous device state information (DSI), and *scheme II* utilizes particle swarm optimization (PSO) technique along with the statistical DSI. The results of simulation on the examination as well as convergence are lastly demonstrated to show the advantages of the proposed distributed KG framework in the fault detection for the electric power systems.

Received on 27 May 2022; accepted on 13 June 2022; published on 14 June 2022

**Keywords:** Big data, knowledge graph, fault detection, electric power systems

Copyright © 2022 Yuzhong Zhou *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eetinis.v9i32.1268

## 1. Introduction

At present, China is in a critical period of integration of industrialization and informatization. Power informatization and intelligence are the inevitable product of "Internet + power" [1–3]. Electric power is the foundation of a country's development [4, 5]. In recent years, smart grid has occupied an important position in the economic field. Due to the wide distribution of smart grid, climate, natural disasters and other reasons, it will lead to power failure, even continuous power failure, resulting in the paralysis of smart grid, which

not only has a serious impact on people's life, but also causes irreparable losses to enterprises and even the whole national economy [6–8]. Therefore, a reliable and accurate power grid fault diagnosis system is of great significance to find fault equipment, diagnose fault causes and eliminate faults in time.

Monitoring system and data acquisition system have long been applied in the early development of power grid. It can feed back the voltage or current change and other electrical quantity information of each device in the power grid in real time through monitoring equipment [9–11], which also provides data guarantee for engineers in power grid fault diagnosis in case of power grid fault [12–14]. However, the power grid is a

\*Corresponding author. Email: [yuzhong\\_zhou@hotmail.com](mailto:yuzhong_zhou@hotmail.com)

dynamic system with complex structure and operation mode, and the causes of faults are diverse. Sometimes the fault signal is not directly related to the causes of faults, which poses a great obstacle to the accurate diagnosis of power grid faults.

The rapid development of big data and artificial intelligence brings opportunities for intelligent diagnosis of power faults. The traditional manual troubleshooting of power grid faults not only consumes a lot of manpower, but also the reliability and accuracy can not be guaranteed, and the location and cause of faults can not be found in time [15–18]. The application of big data mining technology and knowledge graph technology in the field of artificial intelligence to power diagnosis can realize real-time monitoring, prediction and early warning analysis of power grid, shorten the time of troubleshooting, and greatly improve the fault detection accuracy of troubleshooting power grid faults.

By using the newly arising technology of big data and knowledge graph [19–21], this paper analyzes a distributed knowledge graph framework for fault detection in the electric power systems, where multiple devices train their local detection models used for fault detection assisted with a central server. Each device owns its local data set composed of historical fault information and current device state, which can be used to train a local model for fault detection. To enhance the detection performance, the distributed devices interact with each other in the KG framework, where the devices ought to achieve the regional computation in addition to the model aggregation within a specified latency threshold. Through searching for the vibrant qualities together with determined ability at the devices, we enhance the knowledge graph framework by the optimum variety of energetic devices together with the restriction of latency as well as data transmission. Particularly, two data transmission frequency allocation (FA) schemes are developed for the distributed knowledge graph framework, through which *scheme I* is actually bared after the instantaneous device state information (DSI), and *scheme II* utilizes particle swarm optimization (PSO) technique along with the statistical DSI. The results of simulation on the examination as well as convergence are lastly demonstrated to show the advantages of the proposed distributed KG framework in the fault detection for the electric power systems.

The remainder of this particular specific paper is actually orderly as observes. After the intro, Section II discusses the distributed knowledge graph framework under the restriction of the latency for the electric power systems. Afterward, two frequency allocation (FA) methods based upon DSI are actually to enhance the system fault detection accuracy provided in Section III. Additional, simulation outcomes are presented in Section IV.

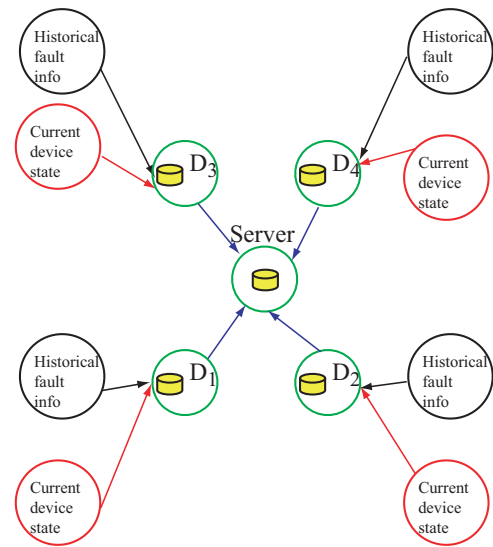


Figure 1. Big data and knowledge graph based fault diagnosis framework for the electric power systems.

## 2. Big data and knowledge graph based fault diagnosis framework

Fig. 1 reveals the big data and knowledge graph based fault diagnosis for electric power systems in the distributed knowledge graph framework, where  $I$  devices  $\{D_i | 1 \leq i \leq I\}$  attempt to speed up their training models for local fault detection, under the assist of a centralized server  $\mathcal{E}$ . Each device  $D_i$  owns a data set  $\mathcal{B}_i$ , composed of historical fault information and current device state. Such data set can be used to train a local model for fault detection, as well as samples in  $\mathcal{B}_i$  is  $|\mathcal{B}_i|$ . The overall number of samples of the fault detection dataset is  $|\mathcal{B}| = \sum_{i=1}^I |\mathcal{B}_i|$ . In reality, the device dataset is constrained, and it is not cozy for each device to achieve an excellent model trained by itself, instead of avoid using other devices' datasets. Whereas, straightly utilizing the datasets of other devices will put a severe burden on the transmission and calculation, leading to a serious problem for information leakage. In order to address these questions, the distributed knowledge graph framework is applied in Fig. 1, in which each device merely considers to transmit the parameters trained by model to the server  $\mathcal{E}$ , which can decrease the transmission energy consumption considerably.

Firstly, the central server  $\mathcal{E}$  arbitrarily chooses  $M$  devices among  $I$  ones for effectiveness as well as justness at every episode [22]. Then, the central server dispatches the whole model (WM) to the all. After that, the device improves the local model (LM) depended on the global model. These  $M$  devices further submit the LMs to the main server  $\mathcal{E}$ , with the aim at speeding up LMs as well as the WM. This procedure is actually repeated  $R$  rounds, unless a penalty deeper system design is actually acquired.

Particularly, at rounded  $r$ , the  $M$ -th device gets the worldwide design of the previous rounded  $\mathbf{v}^{r-1}$ . Generally, the down link data rate is actually much bigger compared to the uplink data rate, as well as the server has actually a larger transfer power compared to the device due to the link along with the energy source. Thus, for device  $D_i$ , the moment to get the worldwide design specification could be minimal. After  $\xi_i$  opportunities of the regional rounds, device  $D_i$  improves the regional setting via utilizing LM design as well as regional dataset, as well as the regional design is actually upgraded as

$$\mathbf{v}_i^t = \mathbf{v}^{t-1} - \rho \nabla A_i(\mathbf{v}^{t-1}), \quad (1)$$

where  $\mathbf{v}_i^t$  represents device  $D_i$  in addition to  $A_i(\bullet)$  means the decrease function for device  $D_i$ . Notation  $\rho$  in addition to  $\nabla$  mean the understanding cost in addition to gradient treatment on the decrease function, particularly. The local informing chance of device  $D_i$  is actually really  $t_i^e$ , offered with

$$t_i^e = \frac{v \xi_i |\mathcal{B}_i|}{a_i}, \quad (2)$$

in which  $v$  is actually the range of CPU cycles computing one instance info in addition to  $\xi_i$  is actually really the epoch of local informing of device  $I$  in a transmission round. Additionally,  $a_i \sim \mathcal{D}(a_{min}, a_{max})$  is actually the computational ability of device  $D_i$ , where  $\mathcal{D}(\bullet)$  means a outfit flow, whereby  $a_{min}$  in addition to  $a_{max}$  are actually really the very little in addition to optimal computational abilities, particularly. After the informing, the local style is actually really sent to the primary internet server, in addition to the transmission chance of device  $D_i$  is actually really offered with

$$t_d^i = \frac{K_i}{Q_i}, \quad (3)$$

where  $K_i$  is actually really the measurement of the style spec of device  $D_i$ , whereby the 32 bits wandering element design is actually really used online, in addition to  $Q_i$  is actually really the transmission cost originating between device  $D_i$  and the internet server  $\mathcal{E}$ , offered with

$$Q_i = O_i \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right), \quad (4)$$

where  $P_i$  is actually really the move power in addition to  $\vartheta_i^2$  is actually really the variance of the additive Gaussian white colored tinted noise (AWGN) at the recipient. The system spec  $g_i$  expertises Rayleigh degree fading together with the common system enhance of  $c_i$ , in addition to  $O_i$  is actually really the cord-free system information move designated to device  $D_i$ . Online, the regularity variety is actually really limited, in addition

to the system information move of the  $I$  devices should please the adhering to restriction,

$$\sum_{i=1}^I O_i \leq O_{total}, \quad (5)$$

where  $O_{total}$  represent the total wireless bandwidth. The internet server  $\mathcal{E}$  extra aggregates the obtained style specs in addition to obtains the around the world style. Our group use the federated common (FedAvg), as well as afterward the around the world style spec  $\mathbf{v}^r$  at pivoted  $R$  might be improved as

$$\mathbf{v}^r \triangleq \frac{\sum_{i=1}^I |\mathcal{B}_i| \mathbf{v}_i^r}{\sum_{i=1}^I |\mathcal{B}_i|}. \quad (6)$$

Coming from the over equations, we can easily draw the conclusion of the overall delay of device  $D_i$  at each  $R$  as

$$t_i = t_i^e + t_i^d. \quad (7)$$

### 3. Optimization of distributed knowledge graph framework

To accelerate the worldwide aggregation, a latency limit  $\beta_{th}$  ought to be actually collection. Particularly, device  $D_i$  has the ability to finish the submit if its own delay is actually listed below  $\beta_{th}$ , i.e.,

$$t_i \leq \beta_{th}. \quad (8)$$

This equation indicates that the device  $i$  will be dropped offline if the latency is above  $\beta_{th}$ . In particular, a smaller  $\beta_{th}$  will increase the probability that the device  $i$  is dropped from the global training. Through integrating this latency restriction, we can easily enhance the distributed knowledge graph framework with allocating the restricted regularity range amongst  $M$  devices, so as to reduce the worldwide reduction work,

$$\min_{\{O_1, \dots, O_M\}} A(\mathbf{v}), \quad (9)$$

subject to  $t_i \leq \beta_{th}, \forall i \in \mathcal{I}$  and  $\sum_{i \in \mathcal{I}} O_i \leq O_{total}$ , where  $\mathcal{I}$  is actually really the compilation of  $I$  devices, in addition to  $A(\bullet)$  represents the around the world decrease function. Since tasks related to data sets may use various decrease functions, the optimization problem in (9) is actually really definitely certainly not fundamental. Affected with that sending more styles efficiently in each pivoted can easily quickly help improve the distributed knowledge graph framework effectiveness [23], our group depend on improve the distributed knowledge graph framework with maximizing the range of energised devices that can easily quickly efficiently send the styles, offered with

$$\max_{\{O_1, \dots, O_M\}} M_1, \quad (10)$$

subject to  $t_i \leq \beta_{th}, \forall i \in \mathcal{I}$  and  $\sum_{i \in \mathcal{I}} O_i \leq O_{total}$ , where  $I_1$  is actually really the range of energised devices, that can easily quickly efficiently send the styles. In the adhering to, we will uses 2 FA strategies in the optimization problem in (10).

### 3.1. Analysis of DSI

As the device state information affects the system performance of distributed knowledge graph significantly, it is of vital important to analyze the DSI associated with the devices, which can be further used for the device fault detection. As to the following DSI,

$$Q_i = O_i \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right), \quad (11)$$

we firstly analyze the expectation, which reflects the long-term value during a long time for the device. The expectation of  $Q_i$  is given by

$$E\{Q_i\} = O_i E \left\{ \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right) \right\}, \quad (12)$$

$$= O_i \int_0^\infty f_{|g_i|^2}(x) \log_2(1 + P_i x / \vartheta_i^2) dx. \quad (13)$$

Note that the PDF of  $|g_i|^2$  is given by

$$f_{|g_i|^2}(x) = \frac{1}{\varepsilon_i} e^{-\frac{x}{\varepsilon_i}}. \quad (14)$$

By applying  $f_{|g_i|^2}(x)$  into (13), we can re-write the expectation of  $Q_i$  as

$$E\{Q_i\} = \frac{O_i}{\varepsilon_i} \int_0^\infty \log_2(1 + P_i x / \vartheta_i^2) e^{-\frac{x}{\varepsilon_i}} dx, \quad (15)$$

$$= -O_i \int_0^\infty \log_2(1 + P_i x / \vartheta_i^2) d e^{-\frac{x}{\varepsilon_i}}. \quad (16)$$

By using the theory of partial integral, we can further obtain the analytical expectation of  $Q_i$  as,

$$E\{Q_i\} = \frac{P_i O_i}{\vartheta_i^2 \ln 10} \int_0^\infty \frac{1}{1 + P_i x / \vartheta_i^2} e^{-\frac{x}{\varepsilon_i}} dx, \quad (17)$$

$$= \frac{O_i}{\ln 10} e^{\frac{\vartheta_i^2}{P_i \varepsilon_i}} \mathbf{Ei} \left( \frac{\vartheta_i^2}{P_i \varepsilon_i} \right), \quad (18)$$

where  $\mathbf{Ei}(\cdot)$  denotes the integral exponential function [24]. Besides the above expectation analysis, we can also analyze the device outage performance, for a given rate  $Q_{th}$ . The outage probability of device  $i$ ,  $P_{out,i}$ , is given by,

$$P_{out,i} = \Pr \left( O_i \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right) < Q_{th} \right) \quad (19)$$

$$= \Pr (|g_i|^2 < \gamma_{th}), \quad (20)$$

where the outage threshold  $\gamma_{th}$  is given by

$$\gamma_{th} = \frac{\vartheta_i^2}{P_i} \left( 2^{\frac{Q_{th}}{O_i}} - 1 \right). \quad (21)$$

By applying the PDF of  $|g_i|^2$ , i.e.,  $\frac{1}{\varepsilon_i} e^{-\frac{x}{\varepsilon_i}}$ , into (20), we can obtain the analytical outage of device  $i$  in the fault detection as,

$$P_{out,i} = 1 - e^{-\frac{\gamma_{th}}{\varepsilon_i}}. \quad (22)$$

### 3.2. I-DSI located FA scheme

As the I-DSI  $M$  devices is actually estimated in addition to acquired at each pivoted of federated understanding, our group can easily quickly refix the FA management problem in (10) according to the immediate DSI. Information that the I-DSI situated FA strategy might be place on the cord-free systems together with repaired fading, like the demand circumstances of repaired Internet of Factors (IoT) bodies. To inform the around the world style a lot a great deal much a lot better, more devices should take part into sending styles in each transmission pivoted. Meant with this, our group recommend a organizing FA strategy, in which the devices together with advantageous system issues have actually the propensity to become really designated an appropriate information move preferentially. Especially, our group to start with type the  $I$  devices in a boiling down acquisition inning conformity along with the system issue, which kinds a compilation  $\mathcal{I}$ . In  $\mathcal{I}$ , the previous devices have really a lot a great deal much a lot better system issue compared with the final. Later, the devices in  $\mathcal{I}$  are actually really selected each other originating from the extremely preliminary to the last, as well as afterward they are actually really designated an appropriate information move, therefore regarding please the latency need. Especially, for device  $D_i$ , its own very personal information move allotment should please the adhering to need,

$$t_i = \frac{K_i}{O_i \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right)} + \frac{v \xi_i |B_i|}{a_i} \leq \beta_{th}, \quad (23)$$

which results in

$$O_i = \frac{K_i}{\left( \gamma - \frac{v \xi_i |B_i|}{a_i} \right) \log_2 \left( 1 + \frac{P_i |g_i|^2}{\vartheta_i^2} \right)}, \quad (24)$$

if certainly there certainly suffices data transfer source left behind. This procedure proceeds up till the devices in  $\mathcal{I}$  have actually been actually totally assigned or even the data transfer source was utilized up.

### 3.3. S-DSI Based FA management

When the over I-DSI located FA scheme needs to understand the I-DSI of all of devices at every rounded, a serious concern is actually enforced on the system application. To reduce this concern, we rely on make use of the stats DSI to refix the FA issue in (10). Details that the S-DSI needs to be actually utilized to the wireless networks, like the request situations of Web of Cars systems. Coming from (23), we first of all compose the network problems along with the provided delay limit  $\beta_{th}$  as

$$|g_i|^2 \geq \left( 2^{O_i \left( \frac{\kappa_i}{\beta_{th} - \frac{v\xi_i|B_i|}{a_i}} \right)} - 1 \right) \frac{\vartheta_i^2}{P_i} = H(a_i), \quad (25)$$

where  $H(y)$  is

$$H(y) = \left( 2^{O_i \left( \frac{\kappa_i}{\beta_{th} - \frac{v\xi_i|B_i|}{y}} \right)} - 1 \right) \frac{\vartheta_i^2}{P_i}.$$

When the networks in the system go through Rayleigh fading,  $|g_i|^2$  observes the rapid circulation along with the typical increase of  $\varsigma_i$ . Within this particular situation, we rely on make the most of the assumption of variety of energetic devices taking part into submitting designs to the server. Intended through this, we first of all determine the possibility that each device can easily effectively submit its own design to the server, which pleases the latency demand. Coming from the possibility thickness work of  $|g_i|^2$ , we can easily acquire the conditional assumption of the  $I$ -th device taking part into submitting its own design to the server as,

$$\begin{aligned} \mathbb{E}(y_i|a_i) &= \Pr\{|g_i|^2 \geq H(a_i)|a_i\} \\ &= \begin{cases} \exp\left(-\frac{H(a_i)}{\varepsilon_i}\right), & a_i > \frac{v\xi_i|B_i|}{\beta_{th}} \\ 0, & a_i \leq \frac{v\xi_i|B_i|}{\beta_{th}} \end{cases}. \end{aligned} \quad (26)$$

Coming from  $\mathbb{E}(y_i|a_i)$  as well as  $a_i \sim \mathcal{U}(a_{min}, a_{max})$ , we could easily compose the assumption  $\mathbb{E}(y_i)$  as

$$\begin{aligned} \mathbb{E}(y_i) &= \int_{a_{min}}^{a_{max}} \frac{1}{a_{max} - a_{min}} \Pr\{|g_i|^2 \geq H(a_i)|a_i\} da_i \\ &\approx \sum_{l=1}^L \frac{\sqrt{1 - \zeta_l^2}}{a_{max} - a_{min}} \exp\left(-\frac{H(\eta_l)}{\varsigma_i}\right), \end{aligned} \quad (27)$$

where the Gaussian-Chebyshev (GC) estimation [24] is actually utilized as well as  $l$  is actually a complexity-vs-accuracy trade off specification along with

$$\zeta_l = \cos\left(\frac{(2l-1)\phi}{2L}\right),$$

$$\begin{aligned} \zeta_k &= \frac{a_{max} - \max\left(a_{min}, \frac{v\xi_i|D_i|}{\gamma_{th}}\right)}{2} \phi_k \\ &+ \frac{a_{max} + \max\left(a_{min}, \frac{v\xi_i|D_i|}{\gamma_{th}}\right)}{2}. \end{aligned}$$

Details that the GC estimation could be precise along with a tool worth of  $k$ . Coming from  $\mathbb{E}(y_i)$ , we can easily determine the assumption of the variety of energetic devices that can easily effectively submit design specifications, provided through

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{n=1}^M \mathbb{E}(y_i), \\ &\approx \sum_{n=1}^N \sum_{n=1}^M \frac{\sqrt{1 - \zeta_n^2}}{a_{max} - a_{min}} \exp\left(-\frac{H(\eta_n)}{\varsigma}\right). \end{aligned} \quad (28)$$

The FA scheme is actually after that developed to make the most of the assumption  $\mathbb{E}(Y)$ ,

$$\max_{\{O_1, \dots, O_M\}} \mathbb{E}(Y), \quad (29)$$

subject to  $\sum_{n=1}^M O_i \leq O_{total}$ . When it is actually challenging to straight refix the optimization issue in (29), we rely on utilize the PSO formula to refix the optimization, which is actually a smart formula based upon populace. In the PSO formula, certainly there certainly are actually  $I$  bits in the populace, as well as each bit consists of 2 essential associates of setting as well as speed. we utilize  $\mathbf{p}_j$  as well as  $\mathbf{w}_j$  to stand for the setting as well as speed of bit  $I$ , specifically, from which  $\mathbf{p}_j = \{O_1, O_2, \dots, O_i\}$  offers a possible service of the data transfer allotment issue in (18) as well as  $\mathbf{v}_j = \{\Delta O_1, \Delta O_2, \dots, \Delta O_i\}$  stands for the increment of  $\mathbf{p}_j$ . Right below,  $\Delta O_i$  is actually the increment of  $O_i$  coming from the present version to the following one. Furthermore,  $\mathbf{pb}$  as well as  $\mathbf{gb}$  are actually utilized to denote the very best FA services of bit  $I$  as well as the worldwide bits up till the present version, specifically, which are actually determined due to the physical health and fitness work. Right below, the physical health and fitness work of the PSO is actually defined through  $\mathbb{E}(Y)$ . At version  $J$ , the speed of the  $I$ -th bit,  $w_i^J$ , is actually upgraded through

$$\begin{aligned} w_i^j &= \pi w_i^{j-1} + q_1 \varphi_1 (\mathbf{pb}_i^{j-1} - \mathbf{p}_i^{j-1}) \\ &+ q_2 \varphi_2 (\mathbf{gb}^{j-1} - \mathbf{p}_i^{j-1}), \end{aligned} \quad (30)$$

where  $q_1$  as well as  $q_2$  are actually 2 velocity constants,  $\varphi_1$  as well as  $\varphi_2$  are actually 2 arbitrary variables consistently dispersed in the variety of  $[0, 1]$ , as well as  $\pi$  means the inertia value element. Coming from (30), the setting  $\mathbf{p}_i^j$  is actually upgraded through

$$\mathbf{p}_i^j = \mathbf{p}_i^{j-1} + \mathbf{w}_i^j. \quad (31)$$

The bits need  $j$  opportunities of version to upgrade their own speed as well as setting inning accordance with (30) as well as (31), specifically. After  $l$  epochs, the  $gb$  would certainly be acquired amongst  $Y$  bits. In further, for  $Y$  bits as well as  $l$  epochs in the PSO formula, the connected calculated intricacy has to do with  $\mathcal{O}(l \times j)$ , in which the efficiency of the PSO could be enhanced along with enhanced varieties of bits as well as iterations.

#### 4. Fault Diagnosis Results and Discussions

Within this particular Section, this paper performs some simulations to validate the proposed studies on the big data and knowledge graph based fault diagnosis in the distributed knowledge graph framework, where Python 3.6 is actually utilized as well as the knowing structure is actually PyTorch 1.8.0. Particularly, the overall interaction rounded is actually readied to 500, the overall variety of devices is actually 200, as well as the variety of chosen devices is actually readied to 10. Otherwise defined, the system overall interaction data transfer is actually readied to 50MHz, as well as the networks in the system expertise Rayleigh level fading, where the typical network increase of the  $I$ -device to the server is actually readied to  $c_i = (j + 50)/200$ , instead of reduction of generality. Furthermore, the transfer energy of devices is actually readied to 0.6 W. In additional, the calculated ability of devices is actually consistently dispersed in the variety of  $[2.1 \times 10^6, 3.1 \times 10^6]$  pattern/2nd. Additionally, for the PSO methods including in the S-DSI located FA scheme, the variety of bit in the populace is actually 30, where the variety of iterations is actually twenty. Furthermore, both velocity constants  $q_1$  as well as  $q_2$  are actually each readied to 0.4 as well as  $\pi = 0.5$ . Additionally, two common datasets of MNIST as well as Fashion-MNIST (FMN) are actually utilized to educate the designs to verify the research researches.

Table I-IV show the fault detection accuracy of various FA schemes versus the interaction rounded, where Table I as well as Table II represent FMN along with  $\beta_{th} = 5.1s$  as well as MNIST along with  $\beta_{th} = 3.1s$ . We can easily observe coming from these 2 numbers that the precision as well as reduction of a number of FA schemes end up being convergent along with the enhancing variety of interaction rounded. In additional, the I-DSI located FA strategy is actually above the S-DSI located one, as well as it can easily accomplish practically the exact very same fault detection accuracy as the standard one, because the instant network condition info is actually efficiently made use of to assist enhance the FA procedure.

Table V reveals the impact of the delay limit  $\beta_{th}$  on the fault detection accuracy of the FA schemes, where the data set FMN is actually utilized as well as  $\beta_{th}$  differs coming from twos to 10s. Coming from Table III, we can

easily discover that the efficiencies of the 2 FA schemes as well as UA enhance along with a bigger  $\beta_{th}$ , as the devices can easily effectively submit the designs to the server more simpler. Furthermore, the I-DSI located FA scheme can easily accomplish practically the exact very same fault detection accuracy as the standard one, as the possibility that devices can easily effectively submit design is actually decreased seriously in the reduced Section of  $\beta_{th}$ . These phenomena validate the 2 FA schemes.

Table VI depicts the effect of overall interaction data transfer on the examination precision of the FA schemes, where the data set FMN is actually utilized,  $\beta_{th}$  is actually fives, as well as the data transfer  $O_{total}$  differs. The examination precision of the several strategies enhances along with the enhancing worth of  $O_{total}$ , as the gear box price ends up being bigger as well as appropriately the gear box latency reduces. Furthermore, the I-DSI located FA scheme can easily accomplish practically the exact very same fault detection accuracy. Besides, compared with the I-DSI FA scheme, the S-DSI located one weakens a lot more quickly along with the lower sized data transfer, because it ends up being harder for the devices to finish the design submit because of the reduced gear box information price in the S-DSI located FA scheme. These reasons additional show the fault detection accuracy of the 2 FA schemes.

Table VII-VIII show the effect of latency limit as well as overall interaction data transfer on the fault detection accuracy of various FA schemes, where the data set uses MNIST. Particularly, Table VII is associated with the fault detection accuracy versus the delay limit along with  $O_{total} = 55MHz$ , while Table VIII is associated with the fault detection accuracy versus the interaction data transfer along with  $\beta_{th} = 3s$ . From these two tables, we can easily discover that the fault detection accuracy of both schemes is better than that of the UA one, and the improvement enhances along with a bigger  $O_{total}$  or even  $\beta_{th}$ , as the devices have actually more chances to finish the regional educating as well as design submit. General, the phenomena in Table VII-VIII shows the fault detection accuracy of the proposed distributed knowledge graph framework with the proposed FA schemes.

#### 5. Conclusions

This paper analyzed a distributed knowledge graph framework for fault detection in the electric power systems, where multiple devices trained their local detection models used for fault detection assisted with a central server. Each device owned its local data set composed of historical fault information and current device state, which could be used to train a local model for fault detection. To enhance

**Table 1.** Fault detection accuracy of FA schemes along with FMN and delay limit  $\beta_{th} = 5.1s$ .

Method \ Round	Fault Detection Accuracy					
	0	100	200	300	400	500
I-DSI based FA	0.189	0.770	0.825	0.855	0.860	0.869
S-DSI based FA	0.189	0.750	0.780	0.799	0.819	0.820
UA	0.189	0.725	0.760	0.779	0.789	0.795
Baseline	0.189	0.780	0.827	0.856	0.861	0.870

**Table 2.** Fault detection accuracy of various FA schemes along with FMN and delay limit  $\beta_{th} = 5.1s$ .

Method \ Round	Loss value					
	0	100	200	300	400	500
I-DSI based FA	2.252	0.611	0.502	0.451	0.421	0.375
S-DSI based FA	2.252	0.701	0.625	0.605	0.523	0.451
UA	2.252	0.911	0.752	0.723	0.625	0.603
Baseline	2.252	0.623	0.511	0.451	0.434	0.375

**Table 3.** Fault detection accuracy of various FA schemes along with MNIST and delay limit  $\beta_{th} = 3.1s$ .

Method \ Round	Fault Detection Accuracy					
	0	100	200	300	400	500
I-DSI based FA	0.108	0.967	0.978	0.982	0.984	0.985
S-DSI based FA	0.108	0.751	0.898	0.941	0.957	0.972
UA	0.108	0.736	0.862	0.904	0.932	0.941
Baseline	0.108	0.983	0.987	0.989	0.991	0.992

**Table 4.** Fault detection accuracy of various FA schemes with MNIST and delay limit  $\beta_{th} = 3.1s$ .

Method \ Round	Loss value					
	0	100	200	300	400	500
I-DSI based FA	2.411	0.131	0.128	0.124	0.123	0.122
S-DSI based FA	2.411	1.129	0.412	0.256	0.143	0.137
UA	2.411	1.144	0.636	0.561	0.515	0.504
Baseline	2.411	0.123	0.122	0.121	0.120	0.120

**Table 5.** Impact of latency limit on the FA schemes along with FMN.

Method \ Threshold(s)	Fault Detection Accuracy				
	2	4	6	8	10
I-DSI based FA	0.861	0.862	0.863	0.863	0.864
S-DSI based FA	0.560	0.805	0.847	0.859	0.862
UA	0.112	0.681	0.830	0.851	0.855
Baseline	0.865	0.865	0.865	0.865	0.865

**Table 6.** Effect of overall interaction data transfer on the FA schemes along with Fashion-MNIST.

Method \ Bandwidth(MHz)	Fault Detection Accuracy						
	50	75	100	125	150	175	200
I-DSI based FA	0.863	0.864	0.865	0.865	0.865	0.866	0.867
S-DSI based FA	0.828	0.840	0.860	0.862	0.863	0.864	0.864
UA	0.792	0.822	0.844	0.852	0.861	0.863	0.864
Baseline	0.865	0.865	0.865	0.865	0.865	0.865	0.865

**Table 7.** Effect of latency limit as well as overall interaction data transfer on the FA schemes along with MNIST.

Method \ Threshold(s)	Fault Detection Accuracy						
	2.0	2.5	3.0	3.5	4.0	4.5	5.0
I-DSI based FA	0.981	0.982	0.985	0.987	0.988	0.991	0.991
S-DSI based FA	0.942	0.952	0.960	0.965	0.966	0.979	0.985
UA	0.651	0.852	0.940	0.960	0.962	0.965	0.970
Baseline	0.992	0.992	0.992	0.992	0.992	0.992	0.992

**Table 8.** Effect of latency limit as well as overall interaction data transfer on the FA schemes along with MNIST.

Method \ Threshold(s)	Fault Detection Accuracy								
	20	30	40	50	60	70	80	90	100
I-DSI based FA	0.981	0.982	0.984	0.986	0.989	0.991	0.991	0.991	0.991
S-DSI based FA	0.885	0.951	0.962	0.971	0.976	0.979	0.980	0.981	0.982
UA.	0.313	0.451	0.525	0.871	0.958	0.963	0.978	0.980	0.982
Baseline	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992

the detection performance, the distributed devices interacted with each other in the KG framework, where the devices should achieve the regional computation in addition to the model aggregation within a specified latency threshold. Particularly, two data transmission FA schemes were developed for the distributed knowledge graph framework, through which *scheme I* was actually based after the instantaneous device state information (DSI), and *scheme II* utilized particle swarm optimization (PSO) technique along with the statistical DSI. The results of simulation on the examination as well as convergence were lastly demonstrated to show the advantages of the proposed distributed KG framework in the fault detection for the electric power systems.

**Acknowledgement.** The work in this paper was supported by the NSFC with grant number 61892236.

## References

- [1] S. Tang and L. Chen, "Computational intelligence and deep learning for next-generation edge-enabled industrial IoT," *IEEE Trans. Network Science and Engineering*, vol. PP, no. 99, pp. 1–12, 2022.
- [2] S. Tang, "Dilated convolution based CSI feedback compression for massive MIMO systems," *IEEE Trans. Vehic. Tech.*, vol. PP, no. 99, pp. 1–5, 2021.
- [3] J. Xia, "Secure cache-aided multi-relay networks in the presence of multiple eavesdroppers," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7672–7685, 2019.
- [4] T. Q. Duong, D. V. Huynh, Y. Li, E. Garcia, and K. Sun, "Digital twin-enabled 6g aerial edge computing with ultra-reliable and low-latency communications," in *Proc. 1st International Conference on 6G Networking, 2022*, pp. Paris, France, 1–6.
- [5] Y. Li, D. V. Huynh, T. Do-Duy, E. Garcia-Palacios, and T. Q. Duong, "Unmanned aerial vehicle-aided edge networks with ultra-reliable low-latency communications: A digital twin approach," *IET Sig. Proc.*, vol. 2022, pp. 1–12, 2022, DOI: 10.1049/sil2.12128.



- [6] C. Li, "Dynamic offloading for multiuser Muti-CAP MEC networks: A deep reinforcement learning approach," *IEEE Trans. Vehic. Tech.*, vol. 70, no. 3, pp. 2922–2927, 2021.
- [7] Y. Guo, "Efficient and flexible management for industrial internet of things: A federated learning approach," *Computer Networks*, vol. 192, pp. 1–9, June 2021.
- [8] S. Tang, "Battery-constrained federated edge learning in UAV-enabled IoT for B5G/6G networks," *Phys. Commun.*, vol. 47, no. 101381, pp. 1–9, 2021.
- [9] J. Sun, X. Wang, Y. Fang, X. Tian, M. Zhu, J. Ou, and C. Fan, "Security performance analysis of relay networks based on-shadowed channels with rhis and cees," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [10] X. Deng, S. Zeng, L. Chang, Y. Wang, X. Wu, J. Liang, J. Ou, and C. Fan, "An ant colony optimization-based routing algorithm for load balancing in leo satellite networks," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [11] C. Wang, W. Yu, F. Zhu, J. Ou, C. Fan, J. Ou, and D. Fan, "Uav-aided multiuser mobile edge computing networks with energy harvesting," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [12] J. Chen, Y. Wang, J. Ou, C. Fan, X. Lu, C. Liao, X. Huang, and H. Zhang, "Albrl: Automatic load-balancing architecture based on reinforcement learning in software-defined networking," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [13] C. Ge, Y. Rao, J. Ou, C. Fan, J. Ou, and D. Fan, "Joint offloading design and bandwidth allocation for ris-aided multiuser mec networks," *Physical Communication*, p. 101752, 2022.
- [14] C. Yang, B. Song, Y. Ding, J. Ou, and C. Fan, "Efficient data integrity auditing supporting provable data update for secure cloud storage," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [15] J. Lu, L. Chen, J. Xia, F. Zhu, M. Tang, C. Fan, and J. Ou, "Analytical offloading design for mobile edge computing-based smart internet of vehicle," *EURASIP journal on advances in signal processing*, vol. 2022, no. 1, pp. 1–19, 2022.
- [16] L. Zhang, W. Zhou, J. Xia, C. Gao, F. Zhu, C. Fan, and J. Ou, "Dqn based mobile edge computing for smart internet of vehicle," *EURASIP journal on advances in signal processing*, vol. 2022, no. 1, pp. 1–19, 2022.
- [17] B. Li, S. Yu, J. Su, J. Ou, and D. Fan, "Computation offloading in multi-uav-enhanced mobile edge networks: A deep reinforcement learning approach," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.
- [18] J. Li, "Snr approximation error analysis for relaying-aided mec-iot networks," *Journal of Engineering*, vol. 2022, 2022.
- [19] K. He, "Learning based signal detection for MIMO systems with unknown noise statistics," *IEEE Trans. Commun.*, vol. 69, pp. 3025–3038, 2021.
- [20] X. Lai, "Secure mobile edge computing networks in the presence of multiple eavesdroppers," *IEEE Trans. Commun.*, vol. PP, pp. 1–12, 2021.
- [21] S. Lai, "Intelligent secure mobile edge computing for beyond 5G wireless networks," *Phys. Commun.*, vol. 45, no. 101283, pp. 1–8, 2021.
- [22] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *AISTATS*, vol. 54, 2017, pp. 1273–1282.
- [23] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 3, pp. 2022–2035, 2020.
- [24] M. Abramowitz, I. A. Stegun *et al.*, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. National bureau of standards Washington, DC, 1972, vol. 55.