Anomaly Detection in DAS Based on Blockchain

ZHOU YUYANG

Background

- What's DAS?

Distributed Automation System, is a system with technologies, including sensors, processors, information and communication networks, and switches, through which a utility can **collect, automate, analyze, and optimize data to improve the efficiency of its distribution power system**.

- Why DAS has abnormalities?

DAS is **NOT absolute automated** in practice. It is a system operated by professional operators. It controls the **power distribution station(PDS)** in different geographical areas to distribute the power. Due to the human operation, abnormalities could occur from time to time.

Background

- What the data in PDS looks like?

Timestamp, I1, I2, I3,…In, V1, V2, V3,…, Vm

Where I means the current and V means the voltage in different parts of power distribution station(PDS). We used these data for anomaly detection to find whether the DAS normal or not.

- What will be the consequences of the anomaly?

Anomaly will negatively affect the grid load and transmission efficiency, and even cause safety hazards to users.

Overview

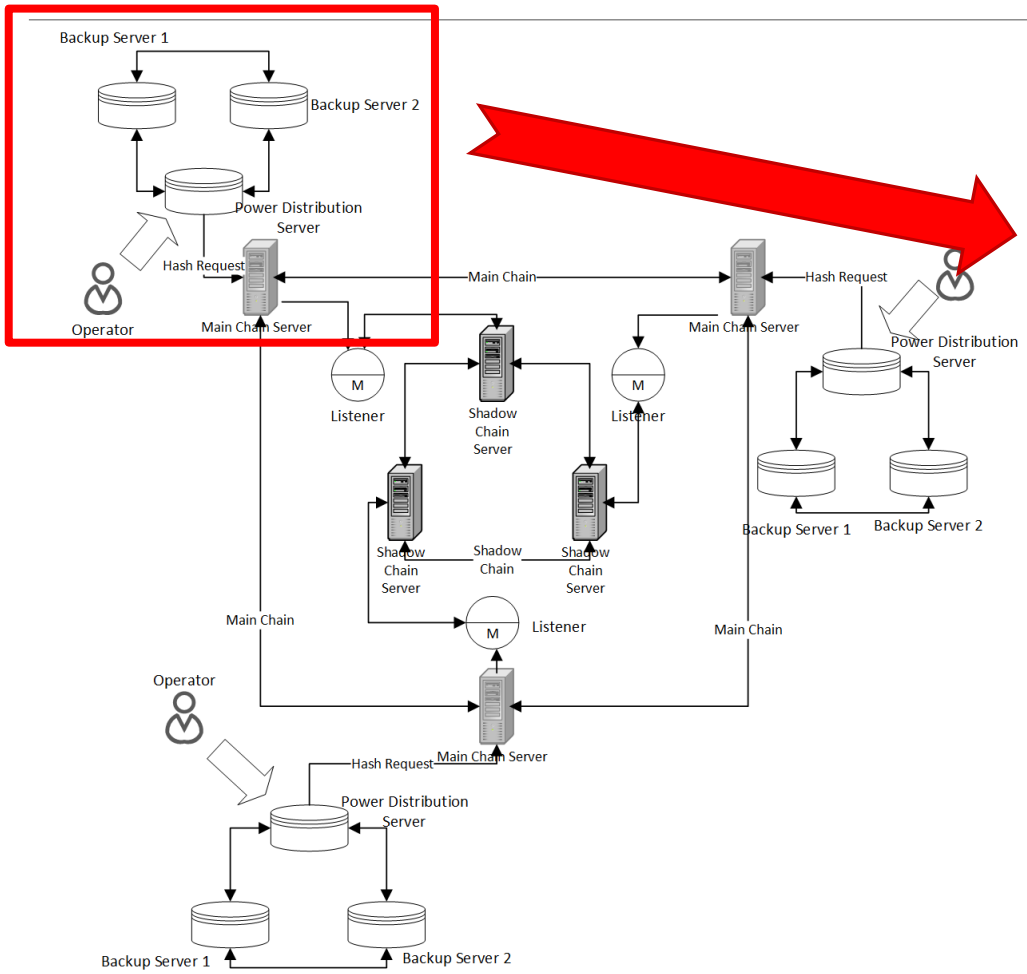
Structure

Double Chain Structure with 2-way Listener

Algorithm

PCA(Principal Components Analysis) and Isolation Forest Algorithm

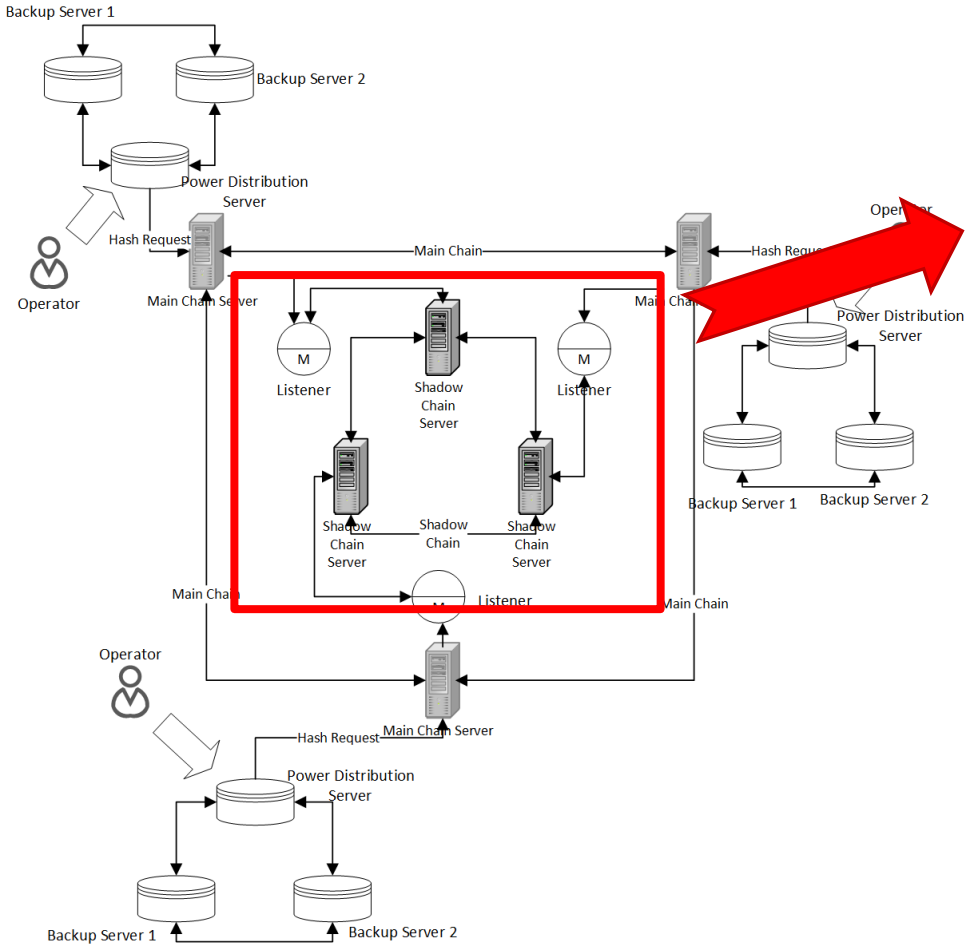
Double Chain Structure



Here's the double chain structure for DAS.

1. **Power Distribution Server** has the following functions: Storage of data and synchronization of data to the two backup databases.
2. **Main Chain Server** accepts the Hash request from the Power Distribution Server and stores the original data and Hash value of the data.
3. **Backup Server** synchronizes information with the Power Distribution Server. When the listener issues a replacement command, the backup server will instead the master power distribution server.

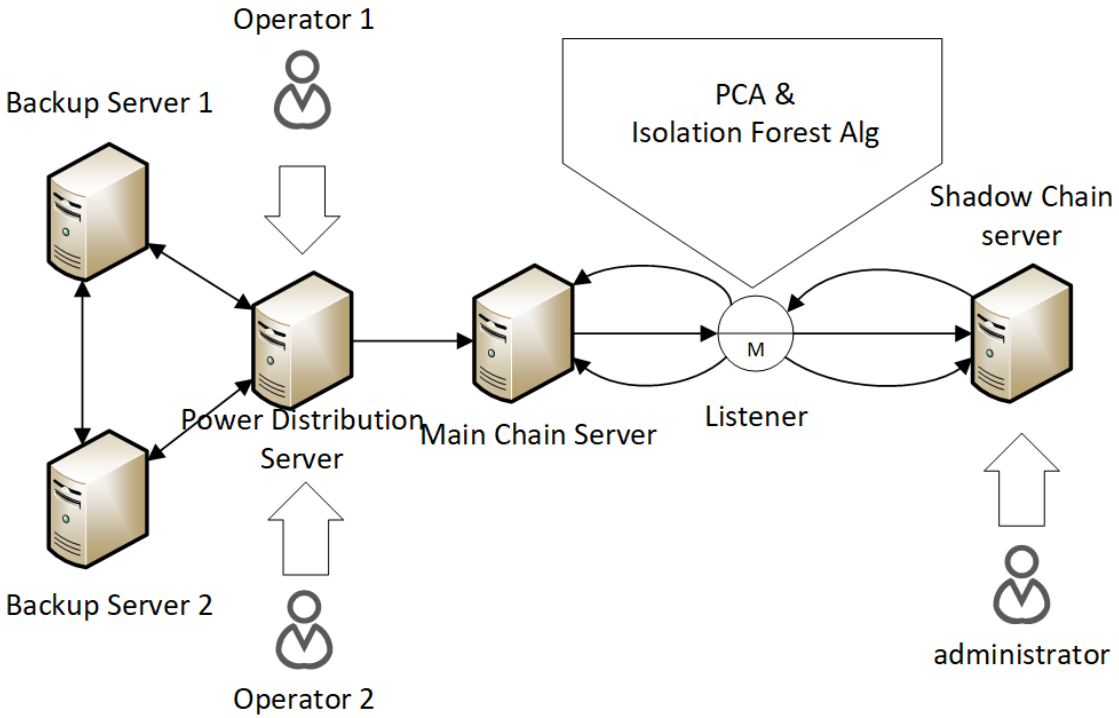
Double Chain Structure



Here's the double chain structure for DAS.

4. The listener is responsible for listening to the main chain server and anomaly detection. If the data is abnormal, the database of power distribution server is rolled back and the abnormal information is recorded in the shadow chain server.
5. If the frequency of data abnormality on a master chain server exceeds the threshold value, the listener removes the server and replaces it by the backup server.
6. Shadow Chain Server records the information of anomaly.

Anomaly Detection Process Chart



Basic Knowledge

- PCA

Principal Components Analysis, is a technique for reducing the dimensionality of large datasets, increasing interpretability but at the same time minimizing information loss.

- Isolation Forest Algorithm

Isolation forest is the World fastest anomaly detection algorithm in terms of algorithmic efficiency. It introduces a fundamentally different method that explicitly isolates anomalies using binary trees, demonstrating the new possibility of a fast anomaly detector that directly targets anomalies without the resource intensive process of normal instance profiling.

PCA & IF

- Isolation Forest Idea

Imagine a scenario where we use a random hyperplane to cut a data space and cut once to generate two subspaces. Next, we continue to randomly select hyperplanes to cut the two subspaces obtained in the first step, and so on, until each subspace contains only one data point. Intuitively, we can see that those clusters with high density are cut many times before they stop cutting, i.e., each point exists in a subspace alone, but those with sparse distribution mostly stop in a subspace very early.

- Alg Process

We randomly generated s pieces of data with n attributes each.

Through PCA, each piece of data is reduced the dimensionalities and has m attributes.

These data point are distributed in a hyperspace. We use isolation forest alg to cut the hyper space and finally got the abnormal data point.

Summary

Aiming at the high cost of distribution data sharing, low transmission efficiency and data security, and the negative impact of abnormal data on power grid load and demand side users during the operation of distribution automation system, an anomaly detection method of distribution automation system based on double chain structure is proposed.

In this method, the distribution data is recorded in the local database and main chain. A two-way listener is added in the main chain server to monitor and detect the system data. If the distribution data is abnormal, the listener commands the distribution station database to rollback and records the abnormal data in the shadow chain server. If data abnormalities occur frequently in a distribution station, replace the data backup to ensure safety.

Anomaly detection algorithm performs principal component analysis on distribution data, and isolated forest algorithm is used for anomaly detection. By simulating the data structure of distribution automation system, generating data and carrying out anomaly detection, the experimental results showed that the algorithm has good applicability to distribution data and can accurately detect abnormal power data.