ATLANTIS PRESS
Advances in Engineering Research, volume 123
2nd International Conference on Materials Science, Machinery and Energy Engineering (MSMEE 2017)

A Cost-effective Parallel Algorithm of Distribution Network Reconfiguration on Hadoop Cluster

Wanxing Sheng [1], Keyan Liu[1], Shupeng Zhou [2], Song Jin [2] and Wei Tang[3]

[1]China Electric Power Research Institute, Beijing 100192, China

[2]Department of Electronic and Communication Engineering North China Electric Power University, Baoding 071003, China

[3]College of Information and Electrical Engineering China Agricultural University, Beijing 100083, China

Abstract: Rapid construction of electric power systems leads to a large increase in quantities of network nodes and data, complicating network reconstruction issue significantly. However, the traditional serial algorithms can not reach a satisfactory computation speed; while some proposed parallel algorithms prior are only applicable to specialized cluster. In this paper, we propose a parallel algorithm of distribution network reconstruction conducted on cost-effective Hadoop cluster. Our algorithm complies with Map-Reduce distributed computing framework. It can process the data of each network node in parallel, thus accelerating power flow calculation. Moreover, our algorithm combines depth-first and breath-first principles together in branch traversals, which substantially improves the probability in finding the optimal solution. Feasibility and effectiveness of the proposed algorithm are verified on a Hadoop cluster.

Keywords: Distribution network; Map-reduce; Parallelization; Power flow calculation; Reconfiguration.

1. Introduction

In recent years, with the attention to the environment, there is a growing hope to reduce the loss of power in the transmission, distribution, and other processes. Distribution network reconfiguration is an important means to achieve the above objectives in distribution system operation and control. However, due to the rapid development of power systems, power networks become more and more complexes, the amount of data needs to be calculated surge. The traditional serial algorithm has been increasingly unable to meet people’s requirements , looking for good parallel method become the new goal.

It is generally believed that cloud computing [1] is a large-scale distributed computing model based on the Internet, which is developed by Distributed Computing, Parallel Computing and Grid Computing. Cloud computing platform based on a machine cluster composed by a large number of cheap PCs , to improve the cost-effective when parallel computing big data. Therefore, using the cloud computing platform to research a parallel network reconstruction algorithm has great significance.

At first, the serial algorithm[2,3] is used to realize the distribution network reconfiguration, such as “Mathematical Optimization Theory Algorithm”, “Switch Exchange Algorithm” and “Artificial Intelligence Algorithm”, etc. With the increasing amount of data, a growing number of people began to study the parallel network reconstruction algorithm and published in the article, for example, “Distribution Network Reconfiguration Based on Parallel Tabu Search Algorithm”, “Application of Undirected Spanning Tree-based Parallel Genetic Algorithm”, “Parallel Genetic Computation for CIM-based Distribution Network Reconfiguration”. However, overall, there is little research on parallel computing for distribution network reconfiguration.

According to the characteristics of a radial distribution network, this paper draws lessons from the relevant knowledge of depth-first traversal and breadth-first traversal，obtains a feasible network topology by this method which based on the random spanning tree search. After that, Newton -

Raphson power flow calculation is transformed into a method with parallel computing capability, and it is applied to the Hadoop platform. By this way, the algorithm not only greatly improves the parallel degree and the calculation speed of big data, but also owns high cost-effective and compatibility for the use of PC cluster.

2. Related Work

Lots of serial algorithms are used to calculate the distribution network reconfiguration. Mathematical Optimization Theory Algorithm[4] can get the global optimal solution which is not dependent on the initial structure of distribution network, But it exists a serious "dimension disaster" problem, and because this algorithm belongs to the greedy-search, the calculation time is long, so it is difficult to meet the actual needs. Switches exchange algorithm[5] was first proposed by S.Civanlar et al. In this algorithm, we started to compute the initial trend and losses， utilizing the results of power flow calculation, the load will be represented by a constant current . Only one contact switch is formed at each time to form a ring network; Then open any one section switch in the ring to make the distribution network recovery to radiation network， so as to achieve the load transfer, to achieve the purpose of load balance and reduce the network loss. This algorithm can quickly determine the distribution network structure which will reduce losses, through heuristic rule cut down optional switch combination, And loss change caused by the switching operation can be estimated via the formula. The eficiency is that you can only operate a pair of switches; It cannot guarantee the global optimum, and the result of the distribution network reconfiguration is related to the initial network structure.

Latterly, the parallel algorithm for distribution network reconfiguration has been greatly developed. “Distribution Network Reconfiguration Based on Parallel Tabu Search Algorithm” [6] introduces divide and conquer. First of all， we choose to the branch group which is disconnected, and use depth-first search algorithm to form a father and son linked list when a refactoring takes place; Then we need to introduce the divide and conquer thought that can reduce the complexity, group by open branch group as the solution space , apply multi-processor for parallel tabu search. This algorithm uses multiple tabu list, effectively expand the search scope, reduces the computational time and increase the efficiency of reconstruction. However, the iterative number of this method is unstable, and it is only applicable to multi-processor systems.

Compared to Undirected Spanning Tree-based Genetic Algorithm，Undirected Spanning Tree-based Parallel Genetic Algorithm introduces the migration operator so that it has the ability to parallel computing. It has excellent stability, better convergence speed, and efficient parallel capability. Nevertheless, its scope of application is relatively small; it is suitable for solving the problem of distribution network reconfiguration with stochastic power flow by using Monte Carlo simulation method, And if you select this algorithm, you must have a specific processor, so its compatibility is not very good.

My algorithm is based on the Hadoop platform, Hadoop is mainly composed of two parts: HDFS and MapReduce . The bottom layer is HDFS, which saves all the files on the storage node of Hadoop cluster, the upper layer is MapReduce . It is a programming model proposed by Google in 2004, which support distributed computing cluster based on massive data, The user through the map interface to deal with a series of <key, value>, and through the reduce interface to calculate each key corresponding to the value. In the open-source cloud computing system, Hadoop ranking first [8].

3. Analysis of Algorithms

3.1 General Overview

In this paper, combined with the characteristics of the distribution network reconfiguration and Hadoop framework, the whole system is divided into four jobs. Its structure is shown below.

Job1’s task is to find the set of adjacent branches. By parallel processing, the first and end nodes of the input data are reassembled according to node number, so that the Job2 can select a set of feasible topology network, Job1 is the foundation of Job2. Job2 is to select a group of branch networks, which is suitable for distribution network reconfiguration and determine its rationality, When selecting the branch, the node number is the order， first of all, we choose a branch connected node 1, then select a branch connected node 2, and so on. The purpose of Job3 is that obtain the nodal admittance matrix. According to the characteristics of the node admittance matrix, we use the MapReduce framework for parallel computing. In the map phase, we take the output data of the key as the line, and parallel compute the value of each row in the reduce phase, if the data is sequentially output in the number of rows, it is a complete node admittance matrix. Job4 iterative compute the node voltage and then calculate the net loss. Since the Jacobian matrix can be block processing, we can divide the matrix with two rows, and calculate the H, N, G and L in parallel.

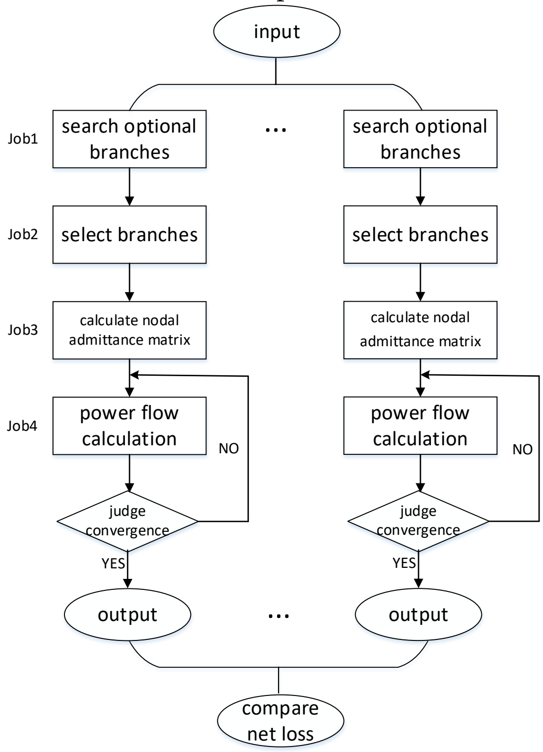

Figure 1 The overall process of our algorithm

We run from job1 to job4 just to get a net loss, if it is a serial program, We need repeatedly cycle procedures, compare their losses, take the minimum. The algorithm of this article based on the Hadoop platform, we can parallel process the overall program which is from job1 to job4， compare net losses after we get all. It can greatly improve the running speed.

$$c_2 = a_2 + b_2 \tag{1}$$

3.2 Select A Feasible Network Topology

Selecting a feasible network topology, that is, to choose a set of branches, this new network, which is composed of them, is still radiate, does not include any loop, and should contain all nodes, Branch =Node—1。

To understand the overall state of the network, we first have to number the branches and nodes，The following diagram show the network topology.

Above is a single feeder six node distribution network, the dotted line indicates link breach. then, we can get the following network topology serial number.

As the newly formed network still be asked radially and containing all the nodes, the traditional depth-first or breath-first traversal is that select a branch first, then, choose the next connected it. The process of generating a new network is shown in Figure 3[9].

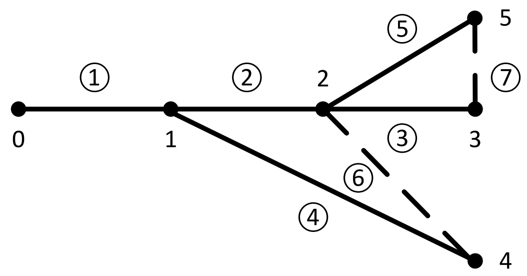

Figure 2 6-node distribution network topology structure

Table 1 network topology serial number

branch number	first node	end node	branch resistance	branch reactance	active load	reactive load
1	0	1	r_1	x_1	P_1	Q_1
2	1	2	r_2	x_2	P_2	Q_2
3	2	3	r_3	x_3	P_3	Q_3
4	1	4	r_4	x_4	P_4	Q_4
5	2	5	r_5	x_5	P_5	Q_5
6	2	4	r_6	x_6		
7	5	3	r_7	x_7		

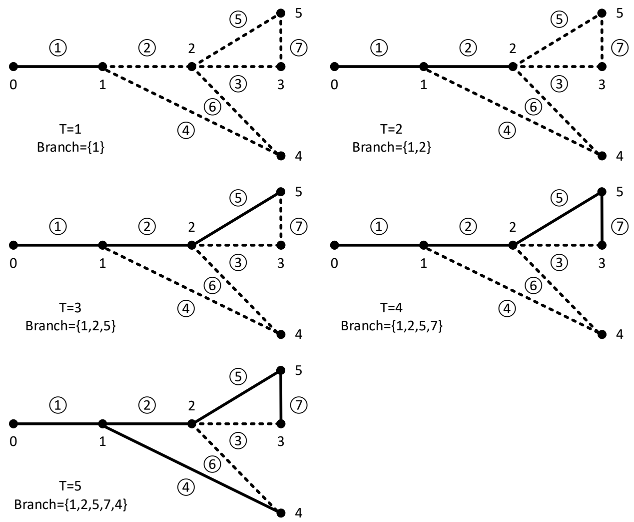

Figure 3 The process of selecting network branches

Assume that the load of a branch is always at the end of the branch, it is not easy to determine which node is the end one during the program run , that will cause difficulties for next Newton - Ralph power flow calculation.

In this paper, we can adjust the order after selecting these branches. Now we should search all branches that are connected to each node. As shown in Figure 4

Node1:①

Node2:②⑤

Node3:⑦

Node4:④

Node5:⑤

Figure 4 All branches connected to each node in the new network

Because the order of the selection is: 1,2,7,4,5, the process of generating new network is shown in Figure 5:

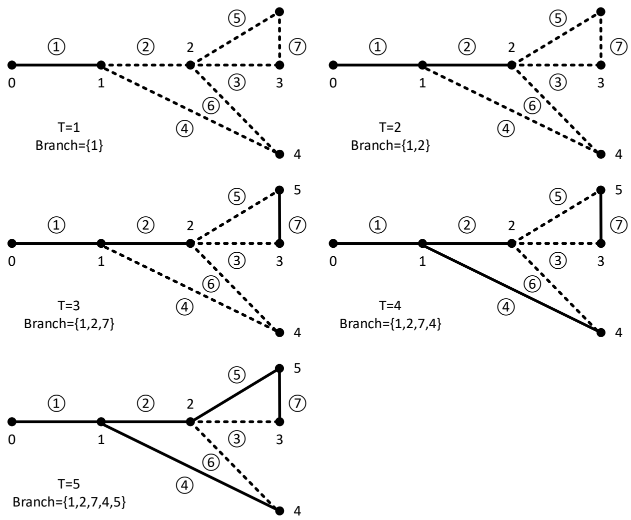

Figure 5 The new process of selecting network branches

The advantage of this method is that the order of the branch is in complete agreement with the sequence of node number. The first selected breach's end node is number one; the second is number two, and so on. It facilitates the subsequent programming calculation.

Job2's work is mainly to select a set of feasible branches, and to determine whether there is a loop, the flow chart is as follows.

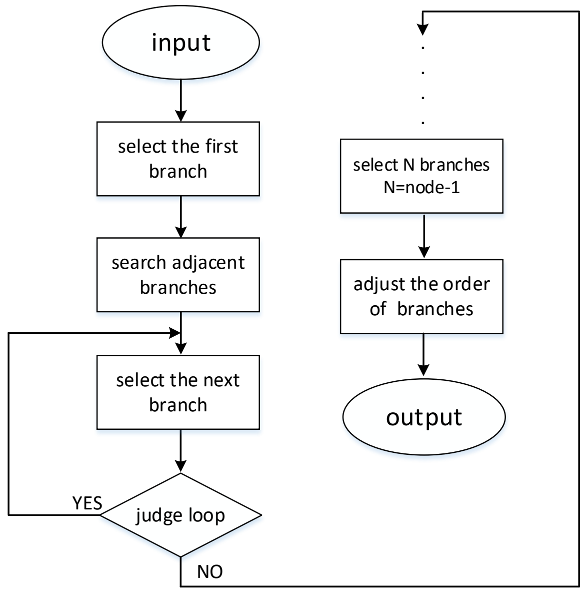

Figure 6 algorithm flow chart of Job2

Step 1: Select the first branch.

Step 2: Random choose one from the next branch group, after, judge whether it has been selected before, if not, carry on with next step, otherwise, re-select one.

Step 3: Determine whether the selected branch group exists loop， if not, carry on with next step, otherwise, return to step 2.

Step 4: Determine whether the next branch group contains one that has never been selected, If not, return to step 1, otherwise, carry on with next step.

Step 5: Repeat steps 2 and 3 until all nodes are traversed. Branch =Node—1.

Through Job1 and Job2, we select a set of feasible network branches and get its all data, then pass it to the next Job.

3.3 Parallel Computing Node Admittance Matrix

As we all know, if we use Newton-Raphson algorithm[10-13] for power flow calculation, we must know the nodal admittance matrix of network structure. The task of Job3 is to compute it in parallel.

By using the data of Table 1, we can get branch resistance and branch reactance, and then solve out the conductance and the susceptance. We set the former as the real part of the mutual admittance, and the latter as the imaginary part .

We assume that Y_{ij} is the mutual admittance of point i and point j , Y_{ii} is the self-admittance of point i , Therefore, the i line of the node admittance matrix represents the admittance of all branches which connected to the node i, if the selection of two nodes are not connected, Y_{ij} equal zero. Due to this characteristic, it is possible to solve the nodal admittance matrix in parallel.

At first,Job3 input the results of Job2, we need to calculate the conductance(R) and the susceptance(X) in the map phase, then, Y_{ij} can be obtained, the output of each map are < First Node， End Node Y_{ij} >, < End Node, First Node Y_{ij} >， In the reduce phase, we take the key as the row, and the first data of the value as the column parallel compute each row of the node admittance matrix. The method shown in Figure 7.

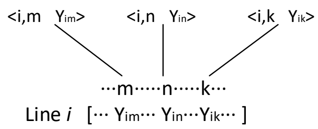

Figure 7 the method for parallel computing node admittance matrix

After getting all the mutual admittances of each line, we can calculated self-admittance according to the formula 1 :

$$Y_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} Y_{ij}$$

Finally, the nodal admittance of each line is sequentially outputed by line number, So we get a complete node admittance matrix on the output file, and pass it to Job4.

3.4 Flow Calculation Based on Newton-Raphson Algorithm.

The mission of Job4 is to compute the power flow by Newton-Raphson algorithm. Since the Job3 has got the node admittance matrix(Y_B), we need to receive the data of Y_B in the back of the table 1 , and transmit the new file to Job4.

Due to the data required by power flow calculations is too much, we need to adjust the output file format of job3, the format as shown in Table 2:

Table 2 The format of Job3's output file

Data Type	N	OP	OQ	G	B	e	f	DP	DQ
Data Number	1	1	1	L	L	L	L	1	1

N is the node number, OP is the active load, and OQ is the reactive load, G represents a real part of the N-th row node admittance matrix, B is the imaginary part, e indicates the real part of all node voltages, and f is the imaginary part, DP as the correction amount of active load, DQ for the reactive load correction, L is the number of nodes.

The procedure that parallel compute Jacobian matrix in map stage is as follows:

（1）Setting the initial value of each node voltage is $e_i^{(0)}$ and $f_i^{(0)}$ .Usually set:

$$U_i^{(0)} = e_i^{(0)} + jf_i^{(0)} = 1.00 + j0$$

（2）Solving the correction amount of the injected power, respectively.

$$DP = P - \sum_{j=1}^{j=n} [e_z(G_j e_j - B_j f_j) + f_z(G_j f_j + B_j e_j)]$$

$$DQ = Q - \sum_{j=1}^{j=n} [f_z(G_j e_j - B_j f_j) - e_z(G_j f_j + B_j e_j)]$$

In the formula, *P* is active load, *Q* for reactive load, *z* is the node number.

（3）Introducing the node injection current

$$I^{(0)} = \frac{P^{(0)} - Q^{(0)}}{\overset{*}{U}^{(0)}} = a^{(0)} + jb^{(0)}$$

（4）Computing Jacobian matrix elements: H、N、G、L.

$$W = \left[\begin{array}{cccccccc|c} H_{11} & N_{11} & H_{12} & N_{12} & \cdots & H_{1n} & N_{1n} & & \\ J_{11} & L_{11} & J_{12} & L_{12} & \cdots & J_{1n} & L_{1n} & & \text{Map1} \\ \hline H_{21} & N_{21} & H_{22} & N_{22} & \cdots & H_{2n} & N_{2n} & & \\ J_{21} & L_{21} & J_{22} & L_{22} & \cdots & J_{2n} & L_{2n} & & \text{Map2} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \end{array}\right]$$

Figure 8 the method for parallel computing Jacobian matrix

This section uses parallel computing, because each line of data is transmitted to the same map, so we can calculate $H_{i1} \cdots H_{in}$ in the i-th map, N 、 J 、 L can also be calculated in the similar. Then they form Jacobian matrix in reduce stage.

In the reduce phase we need to take the inverse of Jacobian matrix, and calculate the variation of each node voltage，finally get the new value of the node voltage. The formula is as follows:

$$e_i^{(1)} = e_i^{(0)} + \Delta e_i^{(0)};$$
$$f_i^{(1)} = f_i^{(0)} + \Delta f_i^{(0)};$$

$\Delta e_i^{(0)}$ and $\Delta f_i^{(0)}$ represent the variation of node voltage.

The new value of the node voltage is used to enter the next iteration, until convergence is reached, then we can calculate the net loss.

The total net loss is ΔS_{Σ} :

$$\Delta S_{\Sigma} = \sum_{i=1}^{i=n} S_i$$

S_i is the power of each node, only the balance node power has unknown.

$$S_1 = U_1 \sum_{j=1}^{j=n} \overset{*}{Y}_{1j} \overset{*}{U}_j$$

S_1 represents the balance node power，$\overset{*}{Y}_{1j}$ is the conjugate value of the element in node admittance matrix，$\overset{*}{U}_j$ indicates conjugate value of each node voltage.

Because it is a single Job iteration, the iterative code written in the main function, so we set Job5 iterate the program of the Job4 .

4. Example Implementation And Result Analysis

4.1 Example Introduction

In this paper, we use the distribution network of IEEE69 nodes[15,16] as experimental example 1, its topological structure is shown in figure.8. Rated Voltage is 12.66kV, Power Reference: 100MVA, Voltage Reference: 12.66kV. We assume that each branch has a switch, the solid line represents a connected branch in the initial state, the dotted line is the branch which is disconnected.

Experimental example 2 come from the literature [17], it contains 118 nodes, the number of contact branches is 15. Due to the network node number is not continuous, does not apply to our program, the number of branch、first node and end node all require rearranging.

The input data of the test system is kept in Text file according to the format of Table 1， and then the program is tested on the Hadoop platform. there are three modes of Hadoop, In this paper, we choose pseudo-distributed mode to emulate.

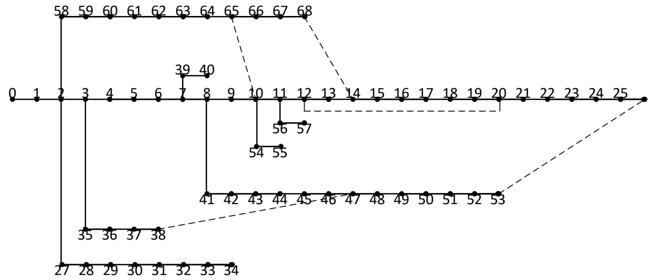

Figure 9 69-node network topology structure

The number of map and reduce in parallel computing is shown in the following table:

Table 3 The number of Map and Reduce

		Job1	Job2	Job3	Job4
Example 1	Map	73	1	68	68
	Reduce	1	1	1	1
Example 2	Map	142	1	117	117
	Reduce	1	1	1	1

4.2 Results and discussion

The input data is actual value, to facilitate the calculation, we need to convert branch resistance, branch reactance, active load and reactive load into Per-unit value(P.u.).

$$Per-unit\ value\left(P.u.\right)=\frac{actual\ value}{base\ value}$$

Table 4 Comparison of example 1 results:

	Original network	Ant colony algorithm	Our algorithm
Contact branch	10-65	10-65	10-65
	14-68	11-12	11-12
	12-20	18-19	18-19
	26-53	51-52	51-52
	38-47	43-44	43-44
Net loss（kW）	225.003	103.36	102.57

Table 5 Comparison of example 2 results:

	Original network	After reconstruction
Net loss（kW）	935.672	813.197

5. Conclusion

The example results show that, the proposed parallel algorithm in this paper can effectively reduce the net loss of distribution network, there is little difference with the result of other methods , it proves the feasibility and effectiveness of this algorithm. When the network contains more nodes and the data is larger, the algorithm's advantage is more obvious. In addition, because of the low cost of the cluster, it is possible to make the commercial promotion. 4. Conclusions

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

References

[1] Tingting Xu. The Power Flow Analysis in Power Systems based on Cloud Computing. MA. thesis, HeFei University of Technology, Anhui, 2012(in Chinese).

[2] Yu Yixin， Qiu Wei， Liu Ruoqin． The distribution network Reconstruction based on heuristic algorithm and genetic algorithm． Power System Technology，35(11):19-22, 2011(in Chinese).

[3] Sarma N D R, Prakasa R K S. A new 0-1 integer programming method of feeder reconfiguration for loss minimization in distribution system. Electric Power Systems Research, 33(2):125-131, 1995.

[4] Zhiqiang Ren,Study on Distribution Netwoek Reconfiguration Based on Improved Genetic Algorithm. MA. Thesis, Northeast Dianli University, Jilin, 2008(in Chinese).

[5] Civanlar S, Grainger J J, Yin H, et al. Distribution Feeder Reconfiguration for Loss Reduction. IEEE Trans on Power Delivery, 3(3):1217-1223, 1998.

[6] Xiaorong Xiang, Dichen Liu, Nong Xiang, Bo Wang. Distribution Network Reconfiguration Based on Parallel Tabu Search Algorithm. Power System Technology, vol.36, pp.100-105, Aug.2012(in Chinese).

[7] Hongcheng Huang, Jie Gu, Chen Fang. Application of Undirected Spanning Tree-based Parallel Genetic Algorithm in Distributed Network Reconfiguration. Automation of Electric Power Systems, vol.39, pp.89-96, Jul.2015(in Chinese).

[8] Chuanmei Wan, Zhenglan Xie. Depth cloud computing:The detailed introduction of Hadoop application development. China Railway Publishing House, 2nd ed , 2014, pp.11-19(in Chinese).

[9] Yannan Ren. Ant Colony System Algorithm For Distribution Network Reconfiguration. MA. Thesis, Xi'an University of Technology, Shanxi, 2007(in Chinese).

[10] Heng Chen. Power system steady-state analysis. china power press, 3nd ed , 2007, pp.129-140(in Chinese).

[11] Ken Chen. Several Approaches to the Method of Newton-Raphson Power-Flow Algorithm in Rectangular Coordinates. Journal of Nanchang University(Engineering & Technology), vol.18, pp.72-77, Jun.1996(in Chinese).

[12] Jun Zhao, Hong Yu, Huasong Zhao. Application of Solving Optimal Power Flow based on Newton-Laphson Method in Electric Power Distribution Network. Agricultural Science & Technology and Equipment, vol.10, pp.34-36, Oct.2014(in Chinese).

[13] Wei Xie, Zhiwei Peng, Chaogang Zhang, Chunsheng Ma. Calculation Method of Power Flow Based on Newton-Raphson. JOURNAL OF XUCHANG UNIVERSITY, vol.25, pp.27-30, Mar.2006(in Chinese).

[14] Boming Zhang, Shousun Chen, Zheng Yan. Advanced Electric Power Network Analysis. Tsinghua University Press, 2nd ed , 2007, pp.24-35(in Chinese).

[15] Lingjun Shi. Study on Distribution Network Reconfiguration Based on Extension Ant Colony Algorithm. MA. Thesis, Nanjing University of Science and Technology, Nanjing, 2009(in Chinese).

[16] Yan Li, Zhiyong Shi, Xiaotao Hai. Distribution Network Reconfiguration Based on Improved Ant Colony Algorithm. YAN JIU YU FEN XI, pp.22-25, Jun.2010(in Chinese).

[17] Dong Zhang, Zhengcai Fu, Liuchun Zhang. An improved TS algorithm for loss-minimum reconfiguration in large-scale distribution systems. Electric Power Systems Research, vol.77, pp.685-694, 2007(in Chinese).