ISSN : 0974 - 7435 Volume 10 Issue 13

2014

BioTechnology
An Indian Journal

FULL PAPER
BTAIJ, 10(13), 2014 [7429-7434]

A topology discovery technique based on computer network

Zhang Chen [1 2 3]
[1]Information Department, Nantong Health College of Jiangsu Province,Jiangsu,226001, (CHINA)
[2]Information Department, Nantong Health Bureau,Jiangsu,226001, (CHINA)
[3]Information Engineer School, Yangzhou University, Jiangsu 225001, (CHINA)

ABSTRACT

This paper presents a new network topology discovery algorithm based on genetic algorithm, and always find the shortest path between any two points, which have strong adaptability. The algorithm and the traditional topology discovery algorithm compared to higher efficiency. At the same time, gives the algorithm model and model-based algorithm. Currently, the algorithm has been applied to the actual network environment.

KEYWORDS

Computer network, Topology discovery, Genetic algorithm.

© Trade Science Inc.

INTRODUCTION

Your goal is to simulate the usual appearance of papers in a Journal of the Academy Publisher. We are requesting that you follow these guidelines as closely as possible.

The main function of Configuration Management in Network Management is the network topology discovery and the network management devices, while the basic of the network management devices are known network devices and its topology[1]. Accurately and timely information on the network topology is the basis for network performance monitoring and evaluation, fault detection and location, resource allocation and management of a series of network management[2,7-8].
In order to establish a network topology discovery mechanism to adapt to changing network environment, people improved the existing algorithm, put forward a lot of new algorithms[3-5,9-10]. But the time complexity of these algorithms generally higher. The advantage of this algorithm is the ability to fully adapt to the high complexity network conditions, and the time complexity is relatively low.

The basic idea of this algorithm is that each node always find the shortest route cost of each other, the route cost is associated with bandwidth, delay, packet loss rate and delay jitter[6]. Each node does not need to know the details of the entire network topology, it only needs to know the route cost to a node. The proposed network topology discovery algorithm, relying on genetic algorithm, with a strong self-adaptability, more effectively compared with traditional algorithms.

ALGORITHM MODEL

Data structures required for the algorithm

List of MAC addresses stored in three

To save space, improve efficiency, the algorithm uses a three list (Figure 1), record the MAC address of the network have visited each node to uniquely identify each node (if there are multiple MAC address, record the smallest one). The first layer as the root, the second layer recording the first 24bit MAC address as network vendor number, the third layer recording the last 24bit MAC address as network number. (Note: To ensure the searching efficiency, the list is defined as an ordered list).

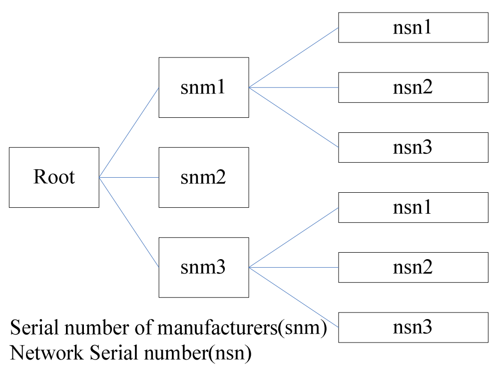

Figure 1 : MAC of the storage structure

Unvisitednodequeue

Use a queue for store node which is waiting for access, the queue only records the n and n +1 layer information. For nodes on the same floor, according to the route cost of the junction to the starting point of the topology discovery from small to large into the queue.
Among, $Node_{n,i+1}$ represents the $Node_{i+1}$ in the current layer and $Node_{n+1,1}$ represents the first node in the next layer. As shown in Figure 2 is currently processing nodes $Node_{n,i}$.

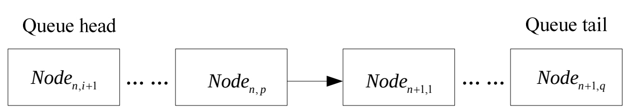

Figure 2 : Processing unvisited node queue

The matrix of rooting cost subnet

In order to improve the efficiency of the algorithm and ensure that every optimal path topology discovery are gone, you need to find the shortest path from a network node to the topology starting point.So, we must find the subnet formed by the pathways from any node to starting point in topology discovery, and record each cost of the pathways. If the two nodes is not the same node and directly connected, we can calculate the cost between which by the function C(x),otherwise the cost between the two nodes is infinite.

3	∞	C(x)	0
2	C(x)	0	C(x)
1	0	C(x)	∞
	1	2	3

Figure 3 : Rooting cost subnet

The Data Structure per Node

In order to preserve their own information and information about each tree node interconnect, we use the following structure.

Node ID	MAC Address	pParent	pChild	pNext

Chart 1 : TheStructure of the Node

The node ID is the node's own identity. If the device has multiple MAC addresses, choose the smallest one. The pParent, pChild, pNext were pointer respectively pointing to the parent node, point to the first child node and point to the sibling node.

Topology Discovery Algorithm in the Model of Computer Network

0. make the gateway of the network management workstation found,as the starting point S,going into the queue UnvisitedNodeQueue,and drawing S.

While (UnvisitedNodeQueue is not empty)

{

a.pop-up a node P from UnvisitedNodeQueue

b.determine the node C (p) directly connected by the following conditions

The port J of the switch Si receives exchanged packets,F (Si,j) is the set of the packets' source MAC addresses. Here is the necessary and sufficient conditions to judge whether the port Xk and the port Sj in two switches connected:

1. F (X,k)∩F (S, j)=null

2. F(X, k)∪F(S, j)=N N is a collection of all nodes.

c f there is no child-node in C_i (p),then continue

d. FOR EACH c IN C(p)

{

(1) get MAC address(Mc) by function GetMAC(c)

(2) if Mc is in list three,continue

(3) calculate the shortest path from c to the starting point(s)

(3.1) get the sub-topology composed of any pathway by function SubTopo(S,c)

(3.2) calculate the routing cost of the adjacent two point V_i and V_j by CalcCost(V_i, V_j),and record the subnet costing matrix corresponding of SubTopo (S, c)

(3.3) get the shortest path from S to c by the genetic algorithm SubTopo (S, c)

suppose V_i is the set of all nodes in SubTopo (S, c)

(3.3.1) Coding

Arrange the nodes in SubTopo (S, c) in a line from top to buttom and left to right. Every node corresponds to a bit.1 indicates a path through a node,and 0 indicates a path does not through a node. In this way, every path can be expressed as follows:

$X = x_1 x_2 x_3 \ldots x_k$ $x_i = 1$or 0

(3.3.2) Fitness evaluation

V_i and V_{i+1} are two adjacent nodes, d(v_i, v_{i+1}) is the routing cost of them. The total routing cost between any two nodes vi and vj in topology D(v_i, v_j)=Sum[d(v_k, v_{k+1})], k=i,i+1,…,j-1. Fitness function is as follow:

Fit (v_i) = C_{max} - D(S, v_i)

if C_{max} > D(S, v_i) otherwise 0

D(S, v_i) is the cost of s and v_i. The smaller the routing cost,the greater the fitness function,the more likely individual genetic to the next group.

(3.3.3)Selecting factor

Selecting factor is “select some individual genetic to the next generation according to certain methods from the father generation group”.

Factoer secletion is according to a certain proportion,based on fitness. the probability of each individual selected increases with the fitness.

M is group esemble, Fiti is the fitness of individual(i). The probability of individual being selected is:

$p_i = Fit_i$ / SUM(Fit_i) i=1,2,3…,M

Then according to the roulette wheel principle,select father individual and genetic to the next generation.

(3.3.4)Cross operation

Cross operation is”two paired chromosome exchange some of its gene in some way, in order to generate two new individual”. Crossover operation is the main way. In the individual code string, set up a cross point randomly,and then exchange some gene of two paired in this point.

If two paired individuals are:

$X1 = x_{11}\ x_{12}\ x_{13}\ldots x_{1j}\ x_{1(j+1)}\ldots x_{1k}\ x_i = 1$or 0

$X2 = x_{21}\ x_{22}\ x_{23}\ldots x_{2j},\ x_{2(j+1)}\ldots x_{2k}\ x_i = 1$ or 0

If the variation is in j, two new individual after variation are:

$X1' = x_{11}\ x_{12}\ x_{13}\ldots x_{1j}\ x_{2(j+1)}\ldots x_{2k}\ x_i = 1$ or 0

$X2' = x_{21}\ x_{22}x_{23}\ldots x_{2j}\ x_{1(j+1)}\ldots x_{1k}\ x_i = 1$ or 0

(3.3.5)Mutation operation

Mutation operation is “some gene mutates into the allele in a group of individual gene,and then generate new individual”. It can improve the ability of local search in algorithm. The basic bit mutation is used here,to detemines everyone whether needs to variation with a certain probability. If it is variatied,we change 0 to 1(1 to 0). The variation of a bit means finding a node.

(3.3.6) Selection of Parameters

Group M=m

Termination of genetic arithmetic T=t

Probability of cross pc = 0.8

Probability of mutation pm = 0.001

}

e.turn the untreated node $C_i(p)$ into queue UnvisitedNodeQueue in accordance with the corresponding shortest path from small to large

f. Adding p to the list stored in three

g. Using Draw (p, C_i (p)),draw C_i (p) and connections

}

APPLICATION OF ALGORITHM

Based on the algorithm,we take a test in the experiment environment

Time	*2014-2-10 15:00*
Place	*Nantong Health Bureau*
Test server	*Two Dell M820*
Test method	*1. Start up two servers, one as topology discovery, another one for drawing.*
	2. Start up SNMP TRAP on all devices in the experiment environment, set the same PUBLIC password.
	3. The first NIC of topology discovery server directly access to the core switching equipment, drawing server connected to the second network card topology discovery server.

Chart 2 : Test environment

Feasibility of the algorithm

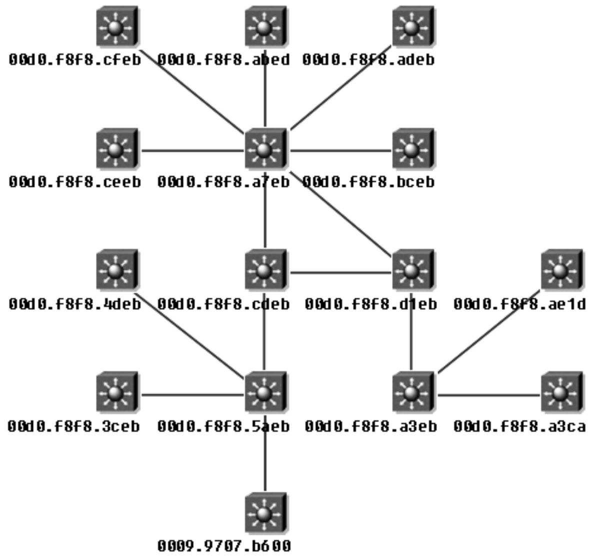

Figure 4 : part of the whole network topology in test environment

Through testing, we found in the test environment topology of the whole network. Here, we show the partial renderings.

Efficiency of the Algorithm

The test results (number of devices found and discovery time) to chart, as follows:

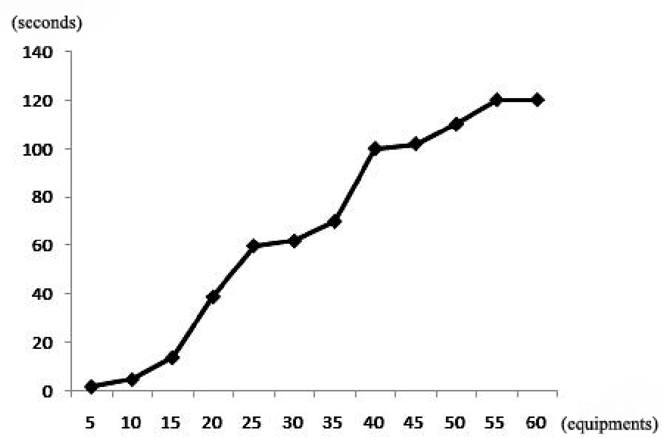

Figure 5 : The relationship between quantity and time equipment found in topology discovery

In this figure the number of devices, found in the initial 10-50 (table), when the elapsed time is relatively long. But the number of devices have been found in 50 (units) above, I found significantly faster. Experimental results show that the larger the network discovery process, the complete algorithm to reflect the accuracy and efficiency.

ALGORITHM FORECAST

Through the SNMP protocol, switch or router to get the relevant Mac and FDB table information, then uses genetic algorithms to select the optimal path discovery network. In this paper, the smallest Mac address as the unique identifier of the network equipment, operation and avoid duplication and redundancy problems accessing the same MAC FDB table and uses subnet spend matrix method, better improve the efficiency of the algorithm, using the Mac three lists, reducing the storage space, it is possible to optimize the time and space complexity of the maximum. Of course, the algorithm is relatively small when the network size, did not achieve a breakthrough on efficiency. This is our future direction of the algorithm improvements.

REFERENCES

[1] Chen Fu, Yang-Jiahai, YangYang; New Algorithms on Ip Network Topology Discovery and its Implement[J], China Journal of Electronics, **8**, 1620-1625 **(2008)**.

[2] Li Xue-feng, Yang Jian-qiang; Research on Physical Topology Discovery Dlgorithm for IP Subnets[J], Computer Engineering and Applications, **45(1)**, 131-133 **(2009)**.

[3] Huang Xiao-bo,Pan xue-zen; The algorithm and realization of network topology discovery[J], Computer Applications and software, **7**, 159-161 **(2007)**.

[4] Zheng Hi, Zhang Guo-qing; An algorithm for physical network topology discovery[J], Computer research and development, **39(3)**, 264-268 **(2002)**.

[5] Qiu Lin, Zhang Jian-zhong,Wu Gong-yi; Research of Physical Network Topology Discovery at Layer 2 Based on Octets [J], Computer Engineering and Applications, **38(22)**, 171-172 **(2002)**.

[6] Yang Yun, Xu Yong-hong, Cao Li-xin, Liu Feng-yu; A New Agent-oriented Distributed Routing Algorithm[J]Computer Engineering and Applications,**12**, 21-22 **(2003)**.

[7] Y.Breitbart, M.Garofalakis, C.Martin, R.Rastogi, S.Seshadri, A.Silberschatz; Topology discovery in heterogeneous IP networks.Proc.of the INFOCOM2000. New York: IEEE Press, 265-274 **(2000)**.

[8] B.Lowekamp, R.David, OpHallaron, Thmoas R. Gross; Topology Discovery for Large Ethernet Network Networks.SIGCOMMp01, **(August2001)**.

[9] Myung-Hee Son, Bheom-Soon Joo, Byung-Chul Ki m, et al.; Physical topology discovery for metro ethernet networks[J].ETRI Journal, **27(4),** 355-366 **(2005)**.

[10] T.Beaubouef, J.Mason; Why the high attrition rate in computer science students : some thoughts and observations[J], SIGCSE Bulletin, **37(2)**, 103-106 **(2005)**.