A Network Topology Discovery Model Based on OSPF

Yaping Jiang[1] ,Peigen Guo[1] ,Mengjia Guo[1]

1. School of Computer and Communication Engineering, Zhengzhou University of Light Industry
Zhengzhou, China
yapingjiang@163.com, peigenguo@163.com, 1205279514@qq.com

Abstract—**Aimed at the shortages of traditional large-scale network topology discovery model in network performance, integrity and real-time performance, when acquiring topology information. According to the characteristics of easy expansion and fast convergence with OSPF protocol, a network topology discovery model OTDM is proposed, based on OSPF. The model adopts passive monitoring to obtain network topology information. After obtaining the link state update (LSU) packet in the OSPF protocol, the Router LSA and Network LSA is analysed to obtain the regional network topology. Then Summary Network LSA is analysed, obtained through boundary routing, to obtain the inter-regional network topology. So a complete network topology is obtained. The experimental results show that the model can accurately and completely acquire large-scale network topology, which can make up for the shortages of existing large-scale network topology discovery technology.**

Keywords—*network topology; OSPF; passive monitoring; link state update packet; Summary Network LSA*

I.   INTRODUCTION

With the development of computer technology, network plays an increasingly important role in the country's politics, economy, military and other fields. Once the network is damaged, almost all social systems will be unable to operate. Network topology can not only enable users to understand the entire network structure, analyze important nodes, data flow and other information in the network, but also efficiently manage the network space, reasonably allocate resources and effectively safety monitor and protect [1].

In recent years, research on large-scale network topology discovery technology has made some progress. Aman Shaikh[2] proposed to use the passive detection technology of OSPF protocol messages to establish the network topology. However, due to the emphasis on analyzing the number and interface of network entities, lack of in-depth analysis of the connection between elements. Spring[3] enabled the Rapid Fuel project to detect the large-scale network topology to enable rapid scanning of network space through distributed, active probing techniques, but a large number of detection packets needed to be sent, which increased the network burden. Gunes[4] proposes an improved algorithm based on OSPF topology detection, which can quickly and accurately analyze and determine various types of complete network topologies, but lacks inter-regional link analysis for large multi-region AS.

In this paper, the large scale network topology discovery model OTDM (OSPF-based topology discovery model) is

proposed to address the shortages of existing large-scale network topology discovery technology. This model can obtain network topology accurately, completely, real-time, and without increasing network burden.

II.   PRINCIPLE OF OTDM MODEL

A large autonomous system is usually divided into multiple areas, which are relatively independent from each other and identified by Area ID [5].The OTDM model uses routers connected to specific OSPF areas to learn the exact same topology data when the network is stable. Each router stores separate LSA data in its own link state database (LSDB).Due to the periodic transmission of LSU carrying LSA, we can get the paper's topological information, such as Area ID, Link State ID, Advertising Router, and Router ID, through analysis by means of passive acquisition, so as to build the network topology by synthesizing the above topological information. The OTDM model proposed in this paper is divided into two modules: one is the regional network topology discovery; the second is the discovery of inter-regional network topology. According to the four network link types in OSPF, the analysis of LSA is better carried out. Fig.1 shows a typical network topology, this paper will use this network topology for analysis.
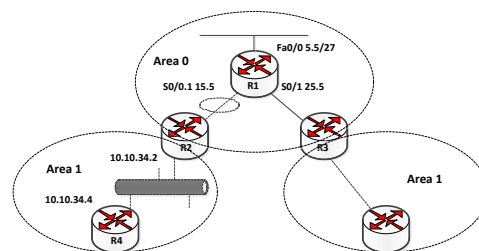

Fig. 1.   Typical network topology

A.   *Regional network topology discovery*

Survivability is a fundamental characteristic of large-scale network, and would not disappear due to system evolution or external environment changes. In the large-scale network, there exists information exchange between the failed subsystems and other subsystems, and thus may cause the cascaded failures. The increase of failed subsystems may cause the whole large-scale network failure.

1)   *Point-to-point and stub network*

It can be seen from Fig.1 that R1 has three stub networks and two point-to-point networks. R2 and R3 also have the same information in the Router LSA. The router has enough information to determine which router to connect to through

the last network, and then uses the IP address configured on the interface to determine the interface connected to other routers. Fig.2 shows the topology types constructed based on the details contained in the R1, R2, and R3 Router LSA.
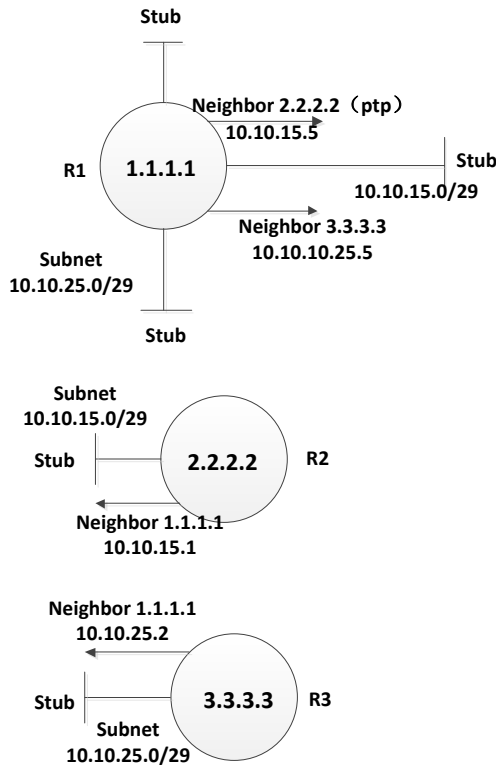

Fig. 2.   Detailed topology type of Router LSA

2) Transport Network

The SPF (Shortest Path First Algorithm) in OSPF requires the use of links between nodes and nodes to build a topology model. Specifically, each link must be between a pair of nodes, but when there are multiple access data links (as shown in Fig.1, the three routers in the left half of area1 are connected in the same subnet), OSPF It must be abstracted to describe the election of the designated router (DR) as a pseudo node. As shown in Fig.3 transmission network.
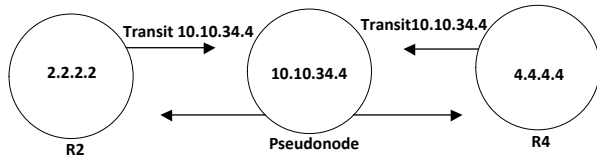

Fig. 3.   Transmission network

3) Virtual Link

The OSPF virtual link allows two ABR connected to the same non-backbone area to establish neighbor relationships through the non-backbone area, even if the two ABR are separated by many other routers and subnets. The virtual circuit appears as a virtual point-to-point link between the two routers.

B.   Inter-regional network topology discovery

The inter-regional network topology is to filter the LSU packets containing the summary network LSA to obtain the area ID, analyze the summary network LSA, obtain the advertising router and the corresponding Link ID field, and obtain all the area ID numbers of an autonomous system from these obtained fields. The ID numbers of the border routers in these areas, as well as all subnets of these areas, can be further analyzed to obtain topology information such as ABR connected to those areas. The inter-regional routing information is transmitted by the type 3 LSA created by the ABR. The inter-regional network topology can be obtained by obtaining the summary network LSA generated by the ABR. One of the important features of OSPF is its flexibility. The protocol can use ABR to connect interconnected networks of any area. ABR can handle a full mesh topology composed of a group of regions (each region is connected to all other regions), a partial mesh topology, a chain topology connecting one region to the next region, or any other network configuration, which can also handle Topological changes that may occur over time.

In the case that it does not understand the internal topology of the area, the ABR only accepts the information passed to it from the area and shares this information with other areas. In Fig.1, the LSU generated by the ABR is captured, and the Area ID and the Summary Network LSA are read. The Advertising router field, thus obtaining the network topology between the regions is shown in Fig.4.
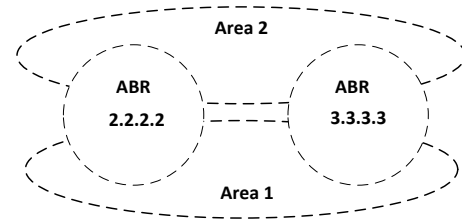

Fig. 4.   Inter-regional network topology

III.   ALGORITHM IMPLEMENTATION

The basic idea of the OTDM model proposed in this paper is to extract the LSU containing the topology information, obtain all the LSA without changing the topology, and perform different operations for the three different types of LSA. After the complete topology list AS is obtained, the algorithm ends. The OTDM algorithm is divided into two steps:

A.   Data processing, establishing a set of regional topology connection relationships.

Defining messages $LSU_i, i = 1, 2, L \ n$, Let V be the node set of a certain area.

$$V = \{v | v = \langle Router\ \mathrm{ID}_i, Area\ \mathrm{ID}_i, AR \rangle\} \qquad (1)$$

$Router\ \mathrm{ID}_i, Area\ \mathrm{ID}_i$, taken from the LSU_i head, $AR \in LSA\ head$.

Let C and N be the collection of connections within the area.

$$C = \left\{ c \mid c = \langle Router\ \text{ID}_i, Area\ \text{ID}_i, AR, \langle Type, Link\ \text{ID}, Link\ Data \rangle \rangle \right\} \quad (2)$$

$Router\ \text{ID}_i$, $Area\ \text{ID}_i$, taken from the LSU_i head, $AR \in Router\ LSA$, the triplet

$\langle Type, Link\ \text{ID}, Link\ Data \rangle$ is abbreviated as T, when there are multiple identical $RouterLSA$, the old information is replaced in the order of reception.

$$N = \left\{ n \mid n = \langle Router\ \text{ID}_i, Area\ \text{ID}_i, AR, Link\ StateID, Attached\ Router \rangle \right\} \quad (3)$$

$Router\ \text{ID}_i\ Area\ \text{ID}_i$, taken from the LSU_i head, $AR \in Network\ LSA$, when there are multiple identical $Network\ LSA$, the old information is replaced in the order of reception.

Connection set C, V, Check that $\exists C.Type \neq trasit$ meets $C.AR = V.AR, C.Link\ ID = N.Link\ State\ ID$ and joins the two sets to get set $CONNET$.

$$CONNET = \left\{ connet \mid connet = \langle Router\ \text{ID}_i, Area\ \text{ID}_i, AR, T \rangle \right\} \quad (4)$$

Connection set C, V, N, check that $\exists C.Type = trasit$ meets $C.AR = V.AR$, $C.Link\ ID = N.Link\ State\ ID$, add the collection connection to $CONNET$.

B. Establish an inter-regional topology connection relationship set, and merge the set to obtain an AS.

Let B be a set of boundary nodes,

$$B = \left\{ b \mid b = \langle Router\ \text{ID}_i, Area\ \text{ID}_{iL\ n}, AR \rangle \right\} \quad (5)$$

$Router\ \text{ID}_i$, $Area\ \text{ID}_{iL\ n}$ from LSU_i head,

$Router\ \text{ID}_i \in \left\{ ABR, ABSR \right\}$, $AR \in Summary\ Network\ LSA\}$

Let E be the set of connections established between the zones,

$$E = \left\{ e \mid e = \langle Router\ \text{ID}_i, AreaID_i L\ AreaID_n, AR, Link\ \text{ID}_i \rangle \right\} \quad (6)$$

Connection set B, E checks if there is $\exists B.Router\ \text{ID}_i \in V.Router\ ID$, satisfies C, and $B.AR\ and\ E.AR$ sets the set $B and E$ to get the set ABR.

$$ABR = \left\{ abr \mid abr = \langle Router\ \text{ID}_i, AreaID_i L\ AreaID_n, AR, Link\ \text{ID}_i \rangle \right\} \quad (7)$$

The CONNET is classified according to the same Area ID to obtain n area sets $CONNET_1 L\ CONNET_n$. Check whether $\exists CONNRT.Type = stub$, and $\exists ABR.RouterID_i \in CONNET.RouterID$ exist in the same Area ID, satisfy $ABR.LinkID = CONNET.LinkID$, and merge the two sets into AS.

$$AS = \left\{ AS \mid RouterID_i, connet_i L\ connet_n, AR \right\} \quad (8)$$

After the above processing, the entire network topology connection relationship set AS is obtained. The above algorithm is used to map the network topology in the area, and then compare and analyze all the network subnets in the obtained area, and then use ABR to connect all the areas to obtain the entire network topology.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

This experiment has set up 20 different detection points throughout the campus to capture OSPF protocol packets flowing through the network. After two weeks of collection, the remaining OSPF packets are only retained by the LSU, and then the OTDM model is used to perform topology discovery on the campus network. It is finally determined that the campus includes one router, 16 three-layer devices, and 145 subnets. Then draw the network topology of the School of Computer and Communication Engineering and the School of Economics and Management. Fig.5 shows the network topology of the two colleges. The area ID is 0.0.0.20, which includes one router, 6 three-layer devices and 33 subnets.
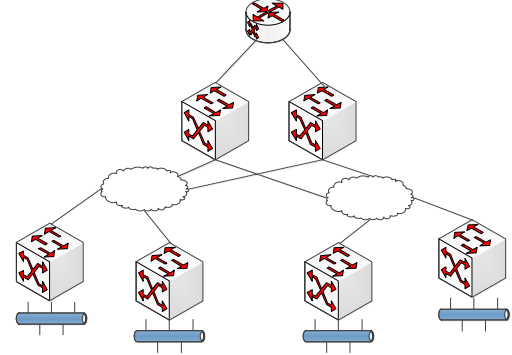

Fig. 5. Two college network typologies

Compared with the actual topology diagram provided by the network center, the number of subnets is less than two, but the remaining link information is basically accurate. The above result will be mainly because the set detection points do not cover all subnets. After obtaining such a conclusion, the area where no detection points have been previously set up is detected, and the topology information of the two subnets that were not previously obtained is obtained. Experiments show that the proposed model can effectively process OSPF packets, and finds a reasonable solution to the difficulties encountered in large-scale network topology discovery in multiple regions.

V. CONCLUSIONS

This paper proposes a large network topology discovery model based on OSPF protocol. Passive detection technology is introduced into the model, which uses layering technology and key bits to obtain a large-scale network topology based on the topology of regional topology detection and inter-regional detection. It avoided the shortages of multi-region network topology, reduce drop detection or false detection, and improve accuracy. However, the OSPF packet acquisition needs to span multiple areas, causing some packets to be incomplete. Combining the collectors to establish tunnels in different areas to obtain complete OSPF Update messages will be the next research direction.

REFERENCES

[1] ZHAO Fan,LUO Xiang-yang,LIU Fen-lin.Research on cyberspace surveying and mapping technology[J].Chinese Journal of Network and Information Security,2016,(09):1-11.

[2] Shaikh A, Greenberg A. OSPF Monitoring: Architecture, Design and Deployment Experience[C]//Proc. of the 1st Symposium on Networked System Design and Implementation. San Francisco, CA:[s.n.], 2004: 57-70.

[3] Spring N, Mahajan R, Wetherall D. Measuring ISP topologies with rocketfuel[J]. ACM Sigcomm Computer Communication Review, 2002, 32(4): 133-145.

[4] Gunes M H, Sarac K. Resolving IP Aliases in Building Traceroute-Based Internet Maps [J]. IEEE/ACM Transactions on Networking, 2009, 17(6):1738-1751.

[5] Charles M.Kozierok.The TCP/IP Guide, A Comprehensive, Illustrated Internet Protocols Reference, published by No Starch Press.