Research on the Anomaly Detection of Flow Streaming Technology in Network

Si Jin[1,a], Sun Guochun[2,b], Zhang Chunhua[3,c]

[1,2,3]Department of Computer, Aviation University of Air Force, Changchun, Jilin, 130022, China

[a]email: sijin_beauty@hotmail.com, [b]email: sungc1963@163.com, [c]email: zch-cc@163.com

Keywords: Network anomaly flow; Flow Streaming Technology; NetFlow

Abstract: According to practical requirements of the network security monitoring, this paper based on flow streaming technology to analyze and research the network anomaly flow. From the perspective of network security, this article introduced the principle of NetFlow technology and the implementation of the collecting network NetFlow streaming, and analyzed the characteristic of anomaly streaming, and then discussed the applications of the collecting network NetFlow streaming technologies in the network anomaly detection.

1 Introduction

In the 21st century, a variety of Internet applications has been rapid popularity. Internet is becoming indispensable information bearing tool in people's daily work life. Enjoying the convenience brought by the Internet, abnormal traffic threat is growing by the Internet at the same time. Abnormal traffic will take up a lot of network bandwidth normal Internet business. In the current network security issues have become increasingly grim situation, how to detect the abnormal traffic behavior quickly, timely, and effectively, this is very important to ensure that the Internet network security obviously.

2 Flow Streaming Technology

Flow is a technical concept which is the network equipment manufacturers in order to improve the routing in the network equipment and the introduction of the internal forwarding speed. The original intention is to transfer from the routing table software query matching operation part of the high CPU consumption to the fast forwarding module of the hardware implementation.

In this mode of function, the packet will be merged into a specific set by several given feature definitions, this set is a Flow. The first data packet of each flow in addition to generate the Flow record, but also drive element three layer module to complete the route query , and put the query results into the Flow record over the same period. The subsequent data packets of the Flow set will be directly in the Flow of the existing records to obtain the route forwarding information. This method can improve the route efficiency of the network equipment.

Flow types are Cflow, sFlow, NetStream and IPFIX usually. With the standardization of IETF, Flow stream data acquisition protocol has been gradually transferred to the NetFlow V9 or IPFIX standards. In this paper, we take V9 NetFlow as an example, to introduce the Flow streaming technology.

2.1 NetFlow V9 information format

NetFlow V9 data portions is divided into three parts: header, template and data. In addition to header, the template and data may have multiple records. The meaning of each part of NetFlow V9 is shown in Table 1.

2.2 Working principle of NetFlow

NetFlow uses the standard exchange mode to process the first IP packet data stream, generates NetFlow buffer. The same data is then transmitted in the same data stream.

NetFlow has two core components. One is NetFlow cache, used to store IP flow information. Another is NetFlow data export or transfer mechanism, used to send data to the network management system. A NetFlow stream is defined as a one-way packet flow between a source IP address and a destination IP address. All data packets have a common transport layer source and destination port number.

Table 1 the meaning of the template and data of NetFlow V9

Name	Position	Length	Meaning
FlowSet ID	Template FlowSet	2 Bytes	Used to distinguish between Template Record and Data Record. FlowSet ID of Template Record is located between 0~255, FlowSet ID of Data Record is more than 255.
Length		2 Bytes	Overall length of the FlowSet.
Template ID	Template Record	2 Bytes	Start a new Template Record, declare a new Data Record format ID, more than 255 in the local network equipment.
Field Count		2 Bytes	Number of fields contained in this Template Record.
Field 1 Type		2 Bytes	Start a new definition of the field, description of the type of the field, Type number related to the manufacturers, 89 types of Cisco are defined in V9 NetFlow.
Field 1 Length		2 Bytes	The length of the field defined above.
……		……	……
Field N Type		2 Bytes	Start a new definition of the field N, description of the type of the field N.
Field N Length		2 Bytes	The length of the field defined above.
FlowSet ID	Data FlowSet	2 Bytes	Refer to a Template Record ID to start a new Data FlowSet. The field values are more than 255.
Length		2 Bytes	Overall length of the Data FlowSet.
Record 1 - Field 1 value	Data Record	2 Bytes	The 1 field values of the first Data Record.
……		……	……
Record 1 - Field N value		2 Bytes	The N field values of the first Data Record.
Record N - Field 1 value		2 Bytes	The 1 field values of the N Data Record.
……		……	……
Record N - Field N value		2 Bytes	The N field values of the N Data Record.

2.3 NetFlow data collection message

Taking NetFlow data of Cisco NetFlow Collector(NFC)as an example, the actual collection data is:

211.*.*.68|202.*.*.195|64917|Others|9|13|4528|135|6|4|192|1

The meaning of each field in the data is: source address | destination address | source domain | destination domain | inflow interface number | outflow interface number | source port | destination port | protocol type | numbers of packages | bytes | stream numbers

In fact, NFC can customize a variety of NetFlow data collection format. Other related manufacturers also provide similar collection software.

3 Anomaly traffic detection system architecture

We need to construct a network anomaly traffic detection system using NetFlow technology. It can be divided into data collection module, data storage module, data analysis module and data display module from the function, system structure as shown in Fig.1. The data display module is output in the form of web pages, and contains all kinds of statistical charts.

3.1 Data collection module

Data collection is one of the key modules. Due to the large NetFlow data traffic, the data can be sampled in time according to the working principle of NetFlow, in order to prevent the network equipment to bring greater data processing pressure. The sampling time interval has a great influence on the performance and the analysis of the data in the later stage. So in the data collection stage, the correct sampling time can be helpful to the more real reducing network environment.

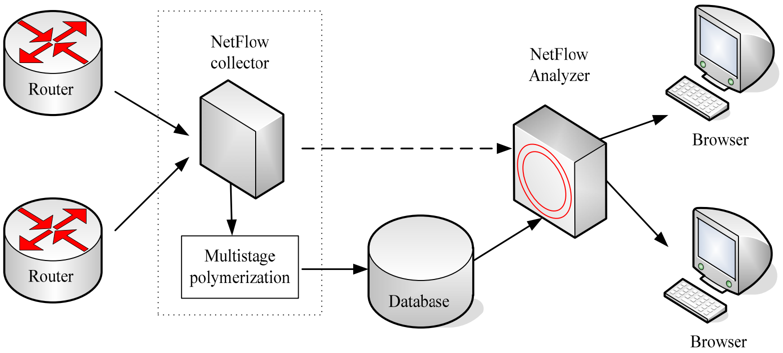

Fig.1 NetFlow network traffic detection system structure

3.2 Data storage module

According to the characteristics of NetFlow data, network traffic analysis can be from two aspects: static and dynamic. Static analysis mainly refers to the NetFlow data in a certain time interval. Dynamic analysis is the analysis of the changing trend of a certain application protocol or a port traffic flow.

3.3 Data analysis module

Static analysis need source/destination host, application protocol, port and other rules for classification of statistics, traffic composition and TopN ranking, etc. Therefore, the corresponding information is extracted from the original flow. According to the original flow information from the UDP package, create the corresponding detailed flow structure, which contains all the data information of the import flow table.

For dynamic analysis, the characteristics of abnormal behavior are important. Abnormal behavior is distinguished as a specific behavior characteristic and no specific behavior characteristic. The dynamic analysis flow chart is shown in Fig.2.

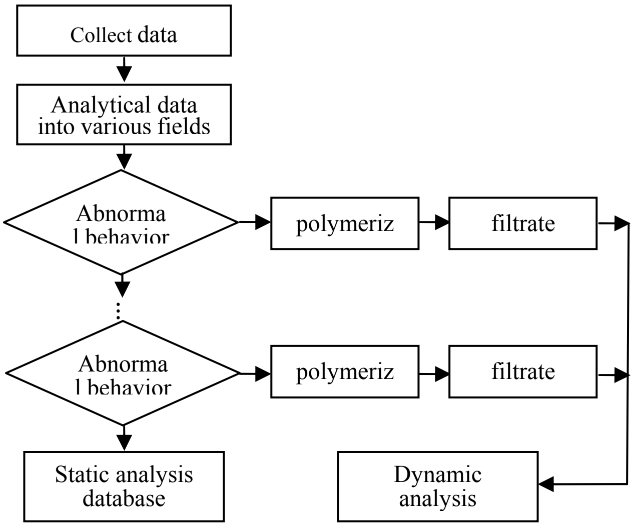

Fig.2 Dynamic analysis flow chart

4 Common abnormal flow analysis

At present, the abnormal traffic of network is mainly caused by network worm traffic, distributed denial of service attack (DDoS) and other network traffic.

4.1 Common worm NetFlow message

For some common worm, it will infect other hosts on the network according to the specific program, so the data packets sent out often have certain characteristics. For example, the protocol type, port number, number of bytes of data packets, which can be judged are their characteristics. The common features of the worm NetFlow packet are as follows:

(1) Code Red Worm

Destination port 80, protocol type 80, packages 3, bytes 144

(2) Worm.opasoft, W32.Opaserv.Worm

Destination port 137, protocol type UDP, bytes 78

(3) WORM.BLASTER, W32.Blaster.Worm

Destination port 135, protocol type TCP, bytes 48

4.2 DDoS message

DDoS has developed a step by DoS to automate this attack. DDoS can coordinate the process on multiple computers to attack. In this case, there will be a flood of denial of service network, which may lead to attack targets due to overload and collapse.

The following is a typical example of the DDoS attack NetFlow data, the case of multiple IP simultaneously to a IP attack:

```
61.*.*.67|69.*.*.100|64821|as9|2|9|49064|5230|17|6571|9856500|1
211.*.*.163|69.*.*.100|64751|as9|3|9|18423|22731|17|906|1359000|1
```

4.3 Other abnormal flow

The other can affect the normal operation of the network can be classified as abnormal traffic. For example, a number of network scanning tool generated by a large number of TCP connection requests, it is easy to make a personality can not be high network equipment paralysis.

The following is an example of NetFlow data, which scan 137 port on the 167.*.211.* network:

```
211.*.*.58|167.*.211.100|65211|as3|2|10|1028|137|17|1|78|1
211.*.*.58|167.*.211.102|65211|as3|2|10|1028|137|17|1|78|1
211.*.*.58|167.*.211.107|65211|as3|2|10|1028|137|17|1|78|1
```

5 Summary

This paper introduces the application of NetFlow technology. We discuss the system architecture of network anomaly traffic detection based on NetFlow from the perspective of abnormal traffic analysis and analyze the characteristics of the data packets of the common abnormal flow, and realize the detection of abnormal traffic, so network management more easy to master the operation of the whole network. It can carry out timely and effective monitoring and early warning of network attacks, viruses and other network abnormal behavior, which provides an effective analysis method for network security management.

References

[1] Guo Jianyun, Cao Qinghua: Research on NetFlow traffic collection and aggregation [J], Modern Electronics Technique, 2009 (7) : 177 - 180.

[2] Pei Wei, Yuan Xiaofang, Wang Dong: Detecting traffic anomalies at application layer in metro network [J], Application Research of Computers, 2010, 27(6)

[3] Xie Tiannian: How to effectively prevent and control network abnormal traffic, China Collective Economy, 2013, 6(2)

[4] Chen Ning, Xu Tongge: Study on NetFlow-based network traffic data collection and storage [J], Application Research of Computers, 2008, 25(2)

[5] Cisco System, Inc. Catalyst 6500 /6000 Switches NetFlow Configuration and Troubleshooting [EB/OL],15 Sep 2006.