Traffic Recognition and Characterization Analysis of MMORPG

Fang Liu, Guotao Yan, Wenli Zhou
School of Information Engineering, Beijing University of Posts and Telecommunications
No.10 Xi Tu Cheng Road, 100876, Beijing, China
lindaliu@bupt.edu.cn

Abstract—A significant share of today's Internet traffic is generated by online games. Unlike traditional services provided by Internet, such as WWW, FTP, SMTP, etc, online games use dynamic ports in transport layer, which causes trouble to traffic recognition from massive Internet traffic. In this paper, a way to recognize MMORPG (Massive Multiplayer online Role Playing Games) traffic is proposed. In this method, MMORPG traffic is recognized by payload inspection. As an illustration of this method, the whole procedure of traffic recognition of Legend in Mir, a popular MMORPG in China, is presented. Furthermore, traffic bandwidth of upload and download, packet length distribution and packet interval distribution are analyzed in detail.

I. INTRODUCTION

With the rapid growth of Internet, more and more traffic is carried over IP network. The online game which is becoming prosperous drastically recent years is a typical example. Reported in a backbone traffic analysis, about 3%-4% of the traffic is attributed to the top 6 popular on-line games[1]. Contrast to other traffic, on-line game has its own characterization, such as sensitive to packets loss and delay, interactivity.

Online Games can be classified into five genres, including role playing games (such as the Legend of Mir), strategy games, fighting games (such as Counter-Strike), management simulation games and entertainment games. The Massive Multiplayer online games (MMORPG) are the most popular online games genre. According to a study in January 2006 in American, 96.2% of the online games market is shared by MMORPG, only 0.9% by fighting games[2]. Although many papers are published on online games, fewer focus on MMORPG, especially popular MMORPG in China. So find an effective way to recognize MMORPG and analyze the traffic characterization is vital for traffic monitoring.

The rest of this paper is organized as follows. In section II, we introduce traffic recognition methods. Section III presents the procedure of MMORPG recognition. Section IV gives the traffic characterization of The Legend of Mir, the popular MMMORPG in China. We arrive at a conclusion at section V.

II. TRAFFIC RECOGNITION METHODS CLASSIFICATION

Historically, the Internet has provided a best-effort service which means that it treated all packets equally in terms of both service quality and pricing. As the Internet is evolving from the best-effort to business quality premium network, a strong demand to measure precise usage of network resources is emerging. Traditionally, Internet traffic was dominated mostly by the client-server type of applications such as WWW, FTP, TELNET, etc. However, this characteristic has been changed significantly when new applications such as peer-to-peer, multimedia and online game applications were introduced. These applications use a range of port numbers or dynamically allocated ones for their sub-transactions. (e.g., EDONKEY uses 4661, 4662, 4665, 6667 and RTSP streaming application allocates a port number dynamically for a stream data transfer). Internet Assigned Numbers Authority (IANA) recommends the usage of application port numbers: 0 - 1023 for well-known ports, 1024 - 49151 for registered ports, and 49152 - 65535 for dynamic and private ports. However, the application developers do not strictly follow this recommendation.

Several Internet applications can use the same port number and some do this for malicious purposes, e.g., port number 80 for security attack. This means that distinguishing flows based on a port number and other header properties is not safe and accurate enough. Thus, application header information and application signature matching in a packet payload are needed for the precise measurement.

Traffic recognition methods can be summarized as the following four types[4].

- Fixed Port-based Recognition Type

Recognition is performed on the basis of a predefined port number to application mapping. Major well known services, such as WWW, FTP, SMTP, BGP, etc., and comparatively popular applications using registered ports can be simply recognized by this method. There exists, however, a rather higher probability of misrecognition due to the current Internet applications characteristics as explained before.

- Payload Inspection-based Recognition Type(Signature Matching)

Recognition is performed on the basis of both port numbers and signatures in the application PDU (Payload Data Unit).

This method produces an effect when two or more equivalently influential applications share a registered port number. Any well known services or popular applications can also be identified by this method if higher level of correctness assurance is required.

- Dynamic Port-based Recognition Type

Recognition is performed on the basis of port numbers obtained by inspecting other flows' payloads. In the sense of payload inspection, this method is similar to Payload Inspection -based Recognition; however, the difference is that, in this type, the sought pattern provides a referential hint to identify another flow that may take place soon after. One of the common examples is a passive mode FTP.

- Reverse Reference-based Recognition Type

Recognition is performed on the basis of a referential note obtained by recognizing a Payload Inspection-based Recognition Type flow on the other links. We define a reverse flow: When there exists a flow, X, of which <src-adr, dst-addr, srcport, dstport, protocol> is <a, b, **x,** y, p>, if another flow, Y, is specified by (b, a, y, **x,** p), then Y is a reverse flow of X. The purpose of the Reverse Reference-based Recognition Type method is, thus, to recognize reverse flows of Payload Inspection-based Recognition Type flows.

III. MMORPG TRAFFIC RECOGNITION

Today, MMORPG with a market share of 96.2% dominates the online game market. In this section we will give an effective way to recognize some popular MMORPG in China. Almost all MMORPG are based on Client/Server mode and use TCP as transfer protocol. As MMORPG server distributed dynamic port, we use Payload Inspection-based Recognition as MMORPG traffic recognition method. We use The Legend of Mir, a popular MMORPG in China as an example to show the traffic recognition procedure.

The Legend of Mir is a popular MMORPG provided by Shanda Company. Most of MMORPG traffic is encrypted, the Legend of Mir is no exception. As the server has to process massive packets from clients and clients are sensitive to delay, the way of encryption can not be too complex.

After payload inspection, we find that the payloads of most packets begins with 0x23 followed by a number (from 1 to 9 in order) as verification and ends with 0x21,except those packets whose payloads contain only one byte (0x2a). The rest of the payload contains game information such as moving, fighting, and so on. The payload, besides the beginning two bytes and the last one byte, is encrypted by BASE64 encryption. The procedure of encryption is as follow: the payload is divided by 3 bytes (24bits) per group. Then each group is divided into four segments. Inserting 2bits (00) at the front of each segment so each segment becomes 1 bytes. At last add 0x3c to each byte and the encryption procedure ends. So after encryption, 3 bytes is expanded 4 bytes, which adds to information redundancy.

Now we focus on payload structure before encryption, besides the first two bytes and the last one byte. The payload format is as follow:

```
DWORD  DW; //(4bytes)
WORD   W1; //(2bytes)
WORD   W2; // (2bytes)
WORD   W3; // (2bytes)
WORD   W4; // (2bytes)
//the above is necessary, they are command codes
CHAR   *charbuffer; //the length is flexible,
//and it is not necessary
```

As command codes, the first 12 bytes is necessary. The character string with flexible length is not necessary. With the encryption arithmetic, we can find out user information from the captured packets. The following is a packet with user information:

```
0040    23 34 3c 3c 3c 3c 3c 42 58 3c 3c 3c 3c 3c
0050  3c 3c 3c 3c 5a 52 59 70 54 63 51 6c 59 3e 79 5e
0060  59 53 3d 70 54 72 74 21
```

The ASCII code after decryption is "#4<<<<g<<<<<<<ygtbupt/buptcn!". The string "ygtbupt" is user account and string "buptcn" is the role name in the game society. Experiential, the packets with user information are logging in packets or packets following by.

With the analysis above, we know that the payload structure and encryption arithmetic. The payloads which start with 0x23 and end with 0x21 is the characteristic of most packets. However, it is vital to find a unique packet in the whole TCP session for the convenience of traffic recognition.

With payload inspection, we find that after logging in, the game client will send a packet with the first nine bytes of 23 33 3c 3c 3c 3c 3c 42 4c(HEX number) and the last byte of 0x21. The length of this unique packet varies with the length of user account.

So we can propose the following method to recognize the Legend of Mir Game traffic: matching signatures packet by packet. If a packet with the first nine bytes of 23 33 3c 3c 3c 3c 3c 42 4c (HEX number) and the last byte of 0x21 is found, take down the source IP, port number, destination IP and port number. Then the packet with the same source IP, port number, destination IP and port number can be recognized as the same flow.

IV. TRAFFIC CHARACTERIZATION ANALYSIS

According to the recognition method above, we can select The Legend of Mir packets from the massive traffic. Analysis shows that the average bandwidth per client is 4Kbps, including 2.5Kbps of downlink (from server to client) bandwidth and 1.5Kbps of uplink bandwidth, which is much lower than the 40Kbps average observed from Counter-Strike[3]. The lower bandwidth is due to the relatively slow motion or active pace in MMORPG. Though the downlink traffic is almost 1.6 times of the uplink traffic, the downlink packets number is only 1.05 times of uplink packets number. So it is obvious that the average length of the packets sent by the server is much bigger than that of the packets from client. Besides, the TCP acknowledgement packets with only heads take a significant part of the whole traffic. In the downlink traffic, the pure ACK packets occupied 34% of the total

packets and 21% of the traffic. In the uplink traffic, the pure ACK packets occupied as much as 55% of the total packets and 50% of the traffic.

Now we focus on the packet size distribution and the interval time between packets from server or client. Note that when we consider packet size, we only consider the length of payload, excluding TCP/IP headers. The pure TCP ACK packets are also not counted. Table I shows the statistic characteristic of packet size while table II shows the statistic characteristic of interval time between packets. In the table, the mode number is the number appears most frequently. Fig.1 to Fig.4. shows the cumulative distribution function of packet size and interval time between packets both from client and server separately.

TABLE I. STATISTIC CHARACTERISTIC OF PACKET SIZE

	Client Packet Size(bytes)	Server Packet Size(bytes)
Mean	16.01	68.42
Median	19.00	18
Mode	19	6863
Std. deviation	6.928	82.848
Variance	47.991	6863
Mininum	1	6
Maxinum	76	524

TABLE II. STATISTIC CHARACTERISTIC OF PACKET INTERVAL TIME

	Client Packet Interval Time(ms)	Server Packet Interval Time (ms)
Mean	395.51	374.65
Median	300	271
Mode	300	0
Std. deviation	477.584	489.804
Variance	228086	239907
Mininum	0	0
Maxinum	18807	18936

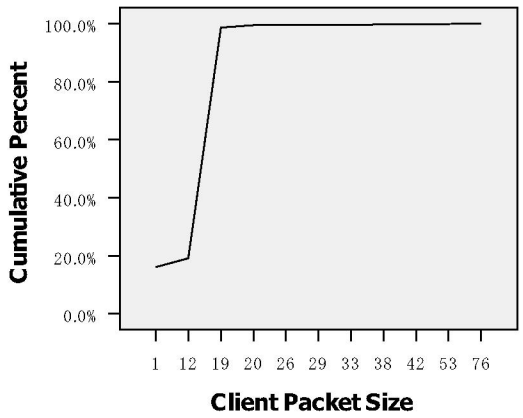

Figure 1. Client Packet Size Distribution

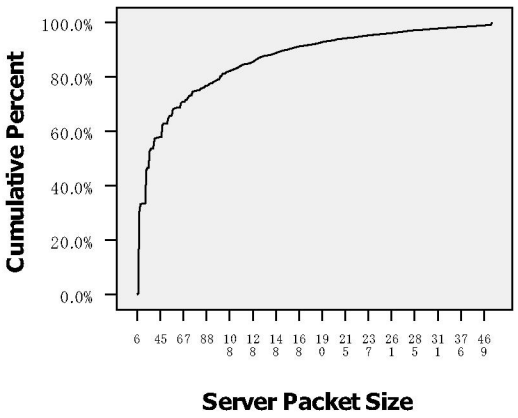

Figure 2. Server Packet Size Distribution

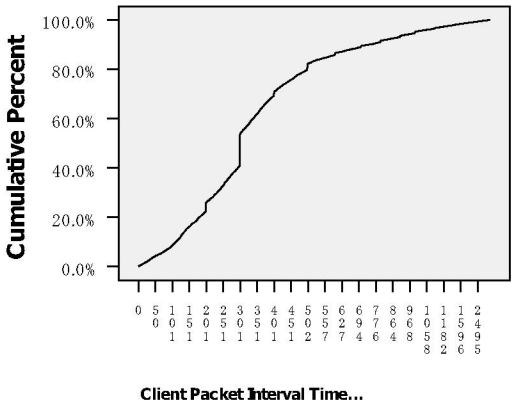

Figure 3. Client Packet Interval Time（ms） Distribution

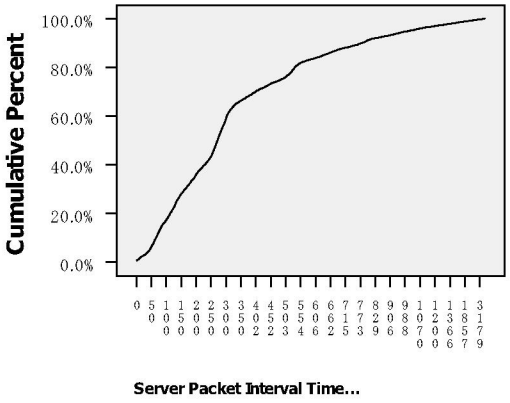

Figure 4. Server Packet Interval Time（ms） Distribution

From the statistic result we know that the average client packet size is far less than the average server packet size. The latter is four times of the former. The packet size of 19 bytes dominates the client packets, with a proportion of over 80%. Less than 2% of packets sent by clients are bigger than 19 bytes. The packets sent by the server have more size types. The packets sent by the server most is the packets with a size of 18 bytes, with a proportion of 30%. The packets whose sizes smaller than 36 bytes occupy half of the total packets.

The interval time between packets both from the client and server differs little. The interval time is 396ms from clients, and 375ms from the server. The client packets with interval time less than 300ms occupies a half, and that less than 500ms occupies 80%, less than 750ms with a proportion of 90%. On the other hand, the server packets with interval time less than 270ms occupies a half, and that less than 535ms occupies 80%, less than 780ms with a proportion of 90%. For the client packet the interval time around 300ms occupies 12.5%. The longest interval time both from the client and server is between 18 seconds to 19 seconds.

Contrast to Counter-Striker, whose interval time between packets is between 50 ms-60 ms[3], the interval time between packets in MMORPG is much longer.

The tiny packets take the dominant part of the online game traffic. So with the rapid growth of online game traffic, we must consider the process of tiny packets when designing network device.

V. CONCLUSION

With the explosion of MMORPG, it is becoming imperative to characterize this component of Internet traffic that will remain sizable portion of overall usage. The MMORPG traffic recognition and characterization analysis is vital to design and optimize network. In this paper we give the way to recognize MMORPG traffic, and the Legend of Mir, a most popular MMORPG in China, is illustrated. At last we summarize the traffic characterization.

REFERENCES

[1] Kuan-Ta Chen, Polly Huang, Chun-ying Huang, Game Traffic Analysis: An MMORPG Perspective, Proceeding of the international workshop on Network and operating systems support for digital and video, June 2005.

[2] http://www.mmogchart.com.

[3] Wu-chang Feng, Francis Chang, Wu-chi Feng. A Traffic Characterization of Popular On-line Games. IEEE/ACM Transactions on Networking, Vol.13, No3. pp.488-500. June 2005.

[4] TS Choi, CH Kim, SH Yoon, JS Park, Content-aware Internet traffic measurement and analysis. Network Operations and Management Symposium, 2004. IEEE/IFIP Volume 1, pp.511 - 524, April 19-23,2004.