A Survey of Techniques for Internet Traffic Identification and Classification

Mingwei Wei, *Member, IEEE*

***Abstract*—The techniques for Internet traffic identification and classification are developed rapidly in recent years, as it widely used in network management, monitor, design, security and research. In the past decade, the traffic identification and classification techniques have been evolved along with development of Internet protocols and applications, and many approaches have been proposed to optimize these techniques. Nowadays, traffic measurement remains one of the hot areas in network research. This is mostly based on the ever increasing network bandwidth, the growth number of network users, the constantly sophisticated applications and the development of technique about confusing traffic identification and classification. In this paper, we present popular traffic identification and classification techniques, include port-based, payload-based, flow-based and host-based, then analyze each technique from challenge aspect and make some remarks and recommendations that contribute to optimize traffic measurement.**

***Index Terms*—traffic identification, traffic classification, challenges, application detection, recommendations.**

I. Introduction

With the development of Internet technology and the advent of the era of mobile Internet, our life has been inseparable from the Internet nowadays. According to the 35th statistical report on development of Internet in China published by CNNIC [2], the Internet users of China have reached 6.49 hundred million by the end of 2014.

As shown in Fig. 1, the Internet population of China has been increasing rapidly in recent years, almost half of Chinese people are using Internet for work or daily life.

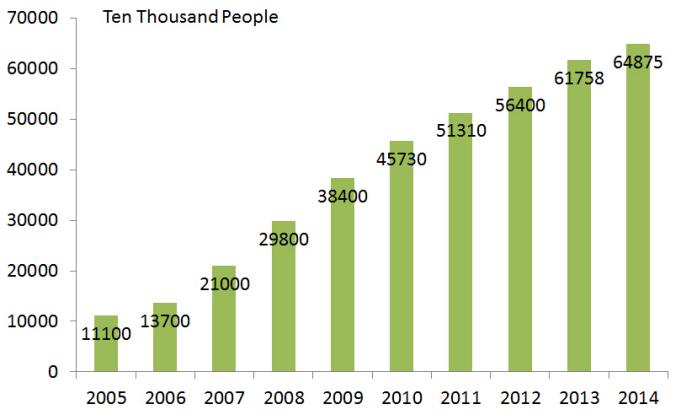

Fig. 1. Internet population of China

Mingwei Wei is with the Department of Computer Science and Technology, Beijing Institute of Technology, Beijing, 100081 China (phone: 152-0161-3264; e-mail: weimw0417@163.com.)

Manuscript received July 25, 2015;

As shown in Fig. 2, the Internet Penetration of China is higher and higher over the years, more and more people felt the charm of the Internet, and the Internet has penetrated into all walks of life of people.

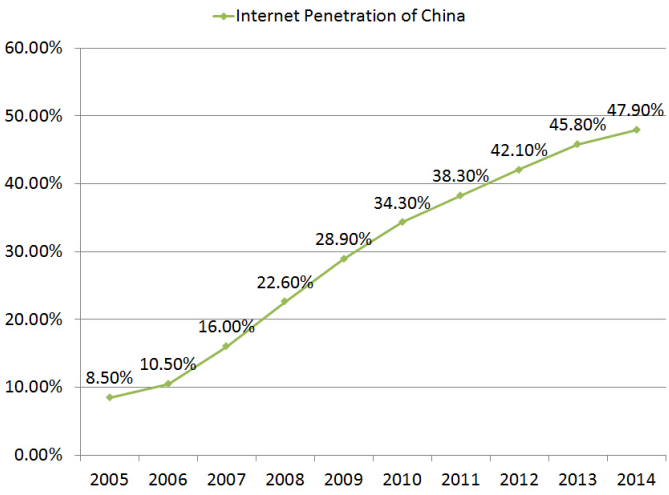

Fig. 2. Internet population of China

Increasingly serious network security problems is in contrast with the rapid development of Internet technology. In the past few years, constantly exposures of network security event make the network security problem get more and more attention. If the security problem is not solved, especially in Internet Finance and Internet Payment, the further development and employment of Internet technology will be severely impeded.

Traffic is the carrier of the Internet. In large number of Internet traffic, a wide variety of malicious traffic is hidden. These malicious traffic carrying viruses, trojans and worms threats the security of Internet. It not only affects the network service provide's service quality, but also threatens the Internet user's privacy and data security, even the national security. So how to find malicious traffic and intercept them is the challenge of Internet security.

Traffic identification and classification is a technique that can detect applications the very traffic corresponded from mass of traffic. Internet traffic identification and classification systems are deployed in gateway normally, it monitors traffic flows though gateway and intercept malicious traffic to ensure smooth operation of network. Traffic identification and classification is basic of traffic control, this technique is widely used in network audit, content audit and intrusion detection, it plays a important role in increasing network management efficiency and guaranteeing network security.

The traditional traffic identification techniques concentrate on content of traffic packet, so it is only to the recognition of unencrypted traffic effectively.

A. Subsection Heading Here

Subsection text here.

1) Subsubsection Heading Here: Subsubsection text here.

II. Conclusion

The conclusion goes here.

Appendix A
Proof of the First Zonklar Equation

Appendix one text goes here.

Appendix B

Appendix two text goes here.

Acknowledgment

The authors would like to thank...

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

[2] CNNIC: Statistics Report of Development of China Internet Network. (Jan 2015). [Online]. Available: http://www.cnnic.cn/hlwfzyj/hlwxzbg /201502/P020150203551802054676.pdf

Mingwei Wei, is the corresponding author of this paper. He is working on master's degree in Network and Information Security Lab in school of computing in Beijing Institute of Technology. His tutor are Mingzhong Wang and Liehuang Zhu. His research interests include network security, business process management and traffic analysis.

John Doe Biography text here.

Jane Doe Biography text here.