THINK INDIA JOURNAL

ISSN:0971-1260
Vol-22- Issue-12-November-2019

An Overview of Various VoIP Traffic Classification Approaches

Max Bhatia (corresponding-author),
Department of Computer Science Engineering,
Lovely Professional University, Punjab (India)
max.16870@lpu.co.in

Vikrant Sharma
Department of Computer Science Engineering,
Lovely Professional University, Punjab (India)
vikrant.23738@lpu.co.in

Abstract: In recent years, various Voice over Internet Protocol (VoIP) applications have emerged which provides real-time transfer of voice communication over the internet. Voice applications such as GTalk, Skype, etc. have gained a lot of popularity because of their ease of use and low cost. However, there exists many reasons such as security, Quality of Service, enforcement of policy, specific interest of business, etc. due to which the monitoring and detection of such traffic is necessary. This paper presents an overview of traditional as well as modern approaches for detecting and classifying VoIP traffic.

Keywords: VoIP traffic, traffic classification techniques

1. Introduction

Voice over Internet Protocol (VoIP) or internet telephony is less expensive alternative for voice communication in comparison to the conventional Public Switched Telephone Network (PSTN). In addition to voice communication, VoIP also supports other services, for example, video calling, messaging, etc. The communication in VoIP occurs in 2 phases: i) signalling and ii) data transmission. Signalling phase establish and disconnect the calls, whereas data transmission phase carries the voice data in the form of IP packets. For the signalling process, there exists various protocols such as H.323, Session Initiation Protocol (SIP), etc. As of late, due to the simplicity and ease of implementation of SIP protocol, it has become the de-facto standard for signalling [1]. There exists standardized applications which support SIP-based telephony [2] or proprietary alternatives like GTalk [4], Skype [3], Yahoo [5], etc. Earlier, VoIP applications were limited to voice communication between computers, but have evolved and has spread to mobile devices like smartphone as well. This has led to the main concern with the telecommunication operators as their revenue from the traditional fixed as well as mobile telephony has reduced. Some organizations along with the telecommunication operators use the mechanism to analyse and identify VoIP traffic for the purpose of charging it or to completely block it. On the other hand, VoIP service providers try to enhance the Quality of Service, reliability as well as security of their applications. Hence, there has been continued attempt to monitor and identify such traffic as new protocols and services emerge. This paper presents an overview of traditional as well as modern approaches which are used to detect and classify VoIP traffic.

The remaining paper is organised as follows: Section 2 presents traditional VoIP traffic classification techniques. Section 3 presents an overview of modern profiling based approaches which include classification on the basis of patterns of network traffic and classification on the basis of communication flow models. At last, conclusion has been presented in Section 4.

ISSN:0971-1260
Vol-22- Issue-12-November-2019

2. Traditional classification techniques

VoIP applications and protocols continue to evolve due to which the traffic analysis and classification process also remains active in the research field. Hence, various detection mechanism have been developed by commercial as well as academic organizations. Most of the techniques are mentioned in [6], [7], [8], [9] describes 2 traditional techniques for classifying VoIP traffic: i) based on the analysis of TCP/IP ports and ii) based on analysis of protocol. These currently these techniques are no longer used as they provide inaccurate results.

i) Port-based analysis: This is the simplest technique to detect the applications as it analyses pre-defined TCP/UDP port numbers which are defined in the packet header. Internet Assigned Numbers Authority (IANA) [10] provides list of TCP/UDP port numbers utilized by various applications and by using this information, organizations could identify the application traffic in their network and could either block them with the help of firewall (or similar mechanisms) or make certain policy over it. But currently, such detection mechanism is inefficient the applications can use the port numbers which are allocated to other protocols (such as HTTP) or use random port numbers for communication. In this way, various applications avoid their traffic from getting detected so that they continue to operate without any hindrance.

ii) Protocol-based analysis: This approach uses hardware/software to collect and monitor the information which is collected from data packets or flows in order to detect the application. But such technique requires prior knowledge about the protocol in the form of application signatures. As new applications emerge or existing ones update their signature patterns, the database of the protocol information needs to be updated frequently. In addition to that, various applications make use of encryption as well, which adds more complexity and hence discourages the used of this technique.

3. Profiling-based approaches

As mentioned in the previous section, traditional classification approaches are no longer efficient since the new applications or protocols keep on emerging which have different behaviour and also hinders the extraction of application signatures or flow model creation or mislead the existing classification techniques is some manner.

a) Classification on the basis of patterns of network traffic

New classification techniques overcomes the issues that existed in traditional classification techniques. There are various research works which provide the mechanism to monitor the behaviour of VoIP applications on the basis of their network usage patterns. A great amount of analysis has been done in [6], [7], [11], [12], [13] which identifies the network characteristics that represents the behaviour of VoIP traffic. Such characteristics are independent of the port-analysis or protocol-analysis and hence overcomes the limitations of traditional classification techniques. The authors in [6] distinguished between the traffic generated by VoIP and non-VoIP applications where they tested Skype, GTalk and Yahoo which have the characteristics of high rate of packets and small packet size in comparison to other applications, for example, chat, file-sharing, games, etc. The studies in [7], [11] distinguished Skype application from other VoIP applications on the basis of average packet

ISSN:0971-1260
Vol-22- Issue-12-November-2019

size, average bandwidth and packet inter-arrival time. An algorithm can be constructed which captures and analyses the behavioural patterns of network traffic generated by various VoIP applications in order to detect them.

It is possible to create an algorithm which captures the network traffic and identify VoIP applications by analysing the network characteristics of VoIP applications or by analysing the behaviour of VoIP applications in the network.

b) Classification on the basis of communication flow models

The approaches mentioned in previous section analyse the specific characteristics which are common in various VoIP applications. However, there exists some VoIP applications which employees tunnelling technique in which the applications hides their traffic behind other applications in order to avoid their detection from firewall or similar mechanisms. The techniques mentioned in previous section are unable to detect such traffic with good amount of accuracy since it introduces noise into the network traffic parameters that are being analysed. Due to this issue, the authors in [8], [9] proposed a technique to analyse the parameters of communication channel where tunnelling is being applied, where they use HTTP protocol by using TCP port 80 or 443. It is possible to create a model from the normal traffic generated by web-browsing which is independent of any traces of VoIP applications. Therefore, by using the statistical characteristics of traffic, HTTP traffic an be profiled, which can be used to train the algorithm for detecting any anomalies with respect to normal web-browsing traffic. In case if anomalies are observed then it can be concluded that some applications are making use of tunnelling technique. The authors in [8], [9] use the parameters: web request size, web response size, number of per-page requests, inter-arrival time between requests and time of page retrieval, for developing an HTTP model and then detected unexpected behaviour in flow characteristics by making use of Chi-square, Goodness-of-fit and Kolmogorov-Smirnov tests.

The detailed comparison, analysis and evaluation of the various VoIP methodologies have been discussed by the authors in [14] where they have compared and summarized various studies with respect to Generic VoIP detection algorithms, identification of Skype traffic on the basis of traffic patterns and analysis of communication flow model. As VoIP application traffic is one of a kind of P2P application traffic, so the detailed discussion regarding various P2P traffic classification approaches along with their comparison has been discussed in [15].

4. Conclusion

Network traffic classification is an important matter for VoIP service providers and telecommunication operators due to various reasons such as: Quality of Service, security, business policies, etc. This paper presents an overview of VoIP traffic classification approaches. Traditional classification techniques include port-based or protocol-based and are inefficient since the current applications try to disguise their traffic or evade detection by using various mechanisms such as using random port numbers, encrypting traffic, etc. Modern VoIP classification techniques falls in 2 categories: i) based on profiling patterns of network traffic and ii) based on flow modelling techniques for anomaly detection. The former approach analyse the network traffic parameters generated by various VoIP applications which are independent of port-based or protocol-based approach. Many algorithms have been in existence which detect the parameters which are common to various VoIP applications or

ISSN:0971-1260
Vol-22- Issue-12-November-2019

target specific type of VoIP application. The latter approach develops a baseline model by analysing the traffic flow of the communication channel and detects any kind of anomaly with respect to the expected behaviour of the traffic flow. Such approach is generally suited for detecting applications which make use of tunnelling technique in order to avoid detection.

References

[1] Dalgic, Ismail, and Hanlin Fang. "Comparison of H. 323 and SIP for IP Telephony Signaling." In *Multimedia Systems and Applications II*, vol. 3845, pp. 106-122. International Society for Optics and Photonics, 1999.

[2] RFC 3261, "SIP: Session Initiation Protocol"

[3] Microsoft Inc., Skype, http://www.skype.com/

[4] Google Inc, Google Talk, http://www.google.com/hangouts/

[5] Yahoo! Inc, Yahoo Message, http://messenger.yahoo.com/

[6] Khan, F. I. U. A. "A generic technique for voice over internet protocol (voip) traffic detection." *IJCSNS* 8, no. 2 (2008): 52.

[7] Perényi, Marcell, and Sándor Molnár. "Enhanced Skype traffic identification." In *Proceedings of the 2nd international conference on Performance evaluation methodologies and tools*, p. 26. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007.

[8] Freire, Emanuel P., Artur Ziviani, and Ronaldo M. Salles. "Detecting VoIP calls hidden in web traffic." *IEEE Transactions on Network and Service Management* 5, no. 4 (2008): 204-214.

[9] Freire, Emanuel P., Artur Ziviani, and Ronaldo M. Salles. "On metrics to distinguish skype flows from http traffic." In *2007 Latin American Network Operations and Management Symposium*, pp. 57-66. IEEE, 2007.

[10] IANA, " Port numbers", http://www.iana.org/assignments/port-numbers

[11] Perényi, Marcell, András Gefferth, Trang Dinh Dang, and Sándor Molnár. "Skype traffic identification." In *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pp. 399-404. IEEE, 2007.

[12] Molnár, Sándor, and Marcell Perényi. "On the identification and analysis of Skype traffic." *International Journal of communication systems* 24, no. 1 (2011): 94-117.

[13] Desclaux, Fabrice. "Skype uncovered–Security study of Skype." *OSSIR–Groupe sécurité Windows* 7 (2005).

[14] Fonseca, Hugo, Tiago Cruz, Paulo Simões, Edmundo Monteiro, José Silva, Pedro Gomes, and Nuno Centeio. "A comparison of classification techniques for detection of VoIP traffic." In *2014 Eighth International Conference on Next Generation Mobile Apps, Services and Technologies*, pp. 117-122. IEEE, 2014.

[15] Bhatia, Max, and Mritunjay Kumar Rai. "Identifying P2P traffic: A survey." *Peer-to-Peer Networking and Applications* 10, no. 5 (2017): 1182-1203.