Identification of encrypted VOIP application using deep learning approach

**KeyWords** deep learning; bi-flow; VOIP; traffic identification

**Abstract** We

VoIP凭借其服务质量以及价格优势获得了越来越多用户的青睐。但因为VoIP服务的便利性也造成了一些社会悲剧。为了使VoIP应用更好的服务于人类，对恶意VoIP的监管是重要的。但是VoIP技术的升级使得传统的识别方法效率低下。

因此，文章针对加密的VoIP应用提出了一种使用deep learning提取语音流量特征的方法，并使用所获得的特征进行voip应用识别。对比现存的VoIP应用的识别方法，本文所提出的方法不需要较长时间的通话语音数据包，因此进行实时的语音识别是可行的。其次，本文提出的提取特征的方法较人为的提取方法更为可靠。

**1. Introduction**

**Recnetly,** VoIP（Voice over Internet Protocol）应用随着其服务质量的提高凭借其优越的价格优势，在广大人群中的得到了越来越多的使用。Voip应用的广泛使用使网络监管的难度大大提高。对voip应用流量进行有效的识别可以帮助解决网络监管问题。

大多数的VOIP应用使用SIP协议或者H.323协议进行通话连接，使用RTP协议进行通话数据传输。随着voip技术的升级，大多数的voip应用会对连接过程和通话传输过程进行加密。在通信建立阶段可能会使用SSL/TLS，SIPS, WEP,WAP/WAP2等加密技术进行加密，在语音传输阶段可能会使用SSL/TLS，SRTP/SRTCP等加密技术进行加密。VOIP服务的完善使得基于端口、特征、模式的识别变得效率低下。

对VoIP应用流量进行较高效的识别是我们亟待解决的问题。如果用户通话时间较长，如何在通话开始阶段进行准确识别；如果通话时间较短，如何保证有足够的特征识别此通话流量。这些问题要求我们需要在较短时间内获取VoIP应用足够多的特征进行特征识别。

深度学习的产生为我们解决以上问题提供了思路，ZhanyiWang提出了将深度学习应用于流量识别领域的思路（The Applications of Deep Learning on Traffic Identification）。深度学习提取的特征较人为提取的特征不但具有更高的可信度，并且也大大提高了识别效率。因此，我们可以对几秒内产生的通话流量进行特征提取并高效识别。本文所提出的方法不但致力于VoIP流量的检测，同时还可以将产生VoIP流量的应用进行分类。

最后，第二节为相关工作介绍；第三节介绍使用深度学习提取应用特征；第四节介绍使用第三节中所提取的特征应用svm进行分类；第五节展示试验结果；第六节总结工作并对下一步工作进行探讨。

**2. Background and Related work**

Khan, F. I. U. A. (2008)提出了一种通用的方法进行VoIP流量检测，文章分析了整个语音流的包间到达时间（packet-inter arrival time）,包大小（packet size），包的交换率（rate of packet exchange）等特征。他们的主要目标是使用通用的特征将VoIP流量与其他流量进行区分，最终得出结论：VoIP流量在每秒内产生的数据包平均个数较其他种类的应用要多（Average packets/Sec rate is greater in VoIP as compared to other applications）；VoIP流量平均包大小要比其他应用要小（Average packet size in bytes is small in VoIP as compared to other applications）。他们的方法在检测VoIP流量的方向上起到了重要意义，但是，该方法并不能针对具体VoIP应用进行分类。

Yildirim, T., & Radcliffe, P. J. (2010, August)提出了一种致力于改善VoIP服务质量，需要较高效的进行VoIP流量识别的方法。文中提出典型的VoIP流量的数据包长度大约在60-150bytes之间。文章主要针对提高VoIP服务质量，所以需要较高效的识别出VoIP流量。但此种方法误差较大，本文也为快速并准确识别VoIP流量提供了思路。

Gomes, J. V., Inácio, P. R., Pereira, M., Freire, M. M., & Monteiro, P. P. (2013)

Alshammari, R., & Zincir-Heywood, A. N. (2015)提出了一种在不包括负载的流量中提取特征集的方法，此方法基于客户端到服务器的双向流，文章使用netmate生成了数据流并对获取的双向流进行特征统计，文中列出了fiat(forward inter-arriaval time), biat(backward inter-arriaval time), fpkt(forward packet length), bpkt(backward packet length), proto, Duration等特征。通过三种有监督的方法，c5.0，Ada Boost和Genetic Programming对训练数据集进行了训练，文章结果显示skype识别率约为100%，其余VoIP应用识别率也在95%左右。文章已取得了较高的准确率，但是识别需要整个数据流的支持。

我们收集了大量的voip流量，本文采用监督学习方法，首先对数据贴上标签。之后按照特定规则进行数据预处理。将其处理成我们所需的形式输入到dnn网络。 得到特征后，使用svm进行分类。对试验结果进行评估。

**3. Methodology**

本文主要对加密的voip应用进行分类操作，学习的过程采用有监督的学习方法。传统的机器学习方法需要人为的去创建特征集，此类方法不但效率低下，并且其准确率没有保障，如采用特征平均包长度来作为特征进行分类，当碰到两个应用的udp报文长度相等时（如本文uucall和altcall）准确率会大大降低。前面提到，本文采用深度学习的思想去获取流量特征，减少人为工作的同时也提升了分类准确率。

**3. Reference**

[1] Wang, Z. (2015). The applications of deep learning on traffic identification. *BlackHat USA*.

[2] Khan, F. I. U. A. (2008). A generic technique for voice over internet protocol (voip) traffic detection. *IJCSNS*, *8*(2), 52.

[3] Alshammari, R., & Zincir-Heywood, A. N. (2015). Identification of VoIP encrypted traffic using a machine learning approach. *Journal of King Saud University-Computer and Information Sciences*, *27*(1), 77-92.

[4] Yildirim, T., & Radcliffe, P. J. (2010, August). VoIP traffic classification in IPSec tunnels. In *Electronics and Information Engineering (ICEIE), 2010 International Conference On* (Vol. 1, pp. V1-151). IEEE.

[5] Gomes, J. V., Inácio, P. R., Pereira, M., Freire, M. M., & Monteiro, P. P. (2013). Identification of peer-to-peer voip sessions using entropy and codec properties. *IEEE Transactions on Parallel and Distributed Systems*, *24*(10), 2004-2014.