# Real Time VoIP Traffic Classification

Lam Hoang Do[1], Philip Branch
Centre for Advanced Internet Architectures, Technical Report 090914A
Swinburne University of Technology
Melbourne, Australia

***Abstract*–This paper presents our research on using machine learning to classify Skype, VoIP and other traffic. We are interested in using a short sliding window rather than the entire flow. We show that the most effective features for classification are packet length and packet interval arrival time. Classifiers constructed to use these features are able to identify traffic into the categories of Skype, VoIP and Other with better than 99% recall when presented with a ten second sample of the flow.**

***Index Terms*–Skype, VoIP, Traffic Classification, Machine Learning**

## I. INTRODUCTION

### A. *Network Traffic Classification:*

The Internet is becoming more and more popular with many new applications and users all over the world. It not only transfers simple data such as text, file and mail but people also share their files using ftp and BitTorrent, watch TV, play online games and communicate with each other using VoIP and Skype. An important question that arises from security issues and quality of service requirements is how network administrators can identify traffic that passes through their networks [1]. There has been some research into identification which has been based on port number and protocol to detect traffic in particular networks. However, this is becoming less reliable with the use of random numbers and Network Address and Port Translation (NAPT). One promising way of identifying traffic flow type is to use machine learning to build up a classifier to identify those traffic by packet statistics such as the maximum packet length, minimum packet length and standard deviation.

In this paper we investigate the use of machine learning techniques to identify traffic generated by a corporate VoIP system using Cisco IP Phones, traffic generated by Skype and other traffic from a busy university backbone network.

### B. *VoIP*

VoIP is the technology that allows voice communication over the Internet Protocol. VoIP systems employ a session control protocol to establish its calls and transmit data. Voice's analogue signal is converted to digital by using an audio codec. The voice payload size depends on the type of Codec we used for the VoIP phone.

In this work we investigate VoIP traffic generated by a Cisco IP Phone system.

### C. *Skype*

Skype is a popular, proprietary Voice over IP (VoIP) application which allows people to make telephone calls over the Internet. Skype users can make phone calls from PC to PC or from a PC to the normal PSTN network.

Skype clients will attempt to use UDP as transport for voice data using a variant of the STUN protocol where necessary, to operate behind a firewall with Network Address translation (NAT). If an attempt to send and receive voice directly via UDP is unsuccessful, both clients which are participating in the call will connect to a Skype Supernode. The Skype Supernode then will work as a relay to enable the conversations between 2 these users [1].

Skype normally uses Sinusoidal Voice Over Packet Coder (SVOPC) while Cisco VoIP uses G.711.

## II. VOIP AND SKYPE IDENTIFICATION ISSUES

### A. *Lawful Interception Obligations*

Governments around the world require telecommunications carriers and manufacturers of telecommunications equipment to provide facilities and services to ensure that they can intercept their systems and services [8].

VoIP and Skype have their own ways of encrypting and protecting the data. It is impossible to detect and collect the content of each call. Fortunately, government agencies are usually only interested in who is communicating with whom rather than the content of the communication. It is only necessary for ISPs and telecommunication companies to provide information about source and destination IP address about traffic that they know it is Skype and VoIP.

### B. *Purpose of Research*

This research is an extension of previous work [1]. That work demonstrated the effectiveness of machine learning in identifying Skype traffic in the presence of other traffic including flows captured on a busy university campus and game traffic. The purpose of this work is to see if the same techniques can be used to classify corporate VoIP traffic (VoIP) and Skype traffic (Skype) in the presence of other non-VoIP traffic (Other).

## III. PROCESS AND RESULTS

### A. *Machine Learning*

Machine learning has been shown to be quite effective at

[1] This author is currently an engineering student at Swinburne University of Technology. This report was written during the author's winter internship at CAIA in 2009

IP traffic classification tasks [1, 12, 13, 14]. Using machine learning to identify particular classes of traffic involves the following steps [3]. First, characteristics of the traffic are identified that might be suitable for classifying the traffic. These characteristics are referred to as features. Features can be associated with single packets, such as packet lengths, or associated with aggregated traffic statistics, such as means and standard deviations. Once the features have been identified the classifier is trained to associate particular features with a particular class of flows. Once trained, the classifier is tested on previously unseen flows. In our work, we are interested in classifying flows into VoIP, Skype and Other classes.

The process of constructing the classifier results in a set of rules or some other mechanism which can then be used to classify previously unseen traffic flows [2, 12, 13, 14].

There are many different techniques for constructing classifiers [2]. The broadest categories are unsupervised and supervised classifiers. Unsupervised classifiers attempt to group objects with similar features into clusters. This kind of classifier is commonly used in identifying different classes that may be present in a particular dataset but is less useful in classifying objects when the classes are already specified.

The other category of classifier is the supervised classifier. Supervised classifiers are presented with examples of object features and the class to which the object belongs. The classifier adjusts its classification mechanism to best match the examples presented to it. It 'learns' that certain combinations of features are associated with certain classes. There are a number of different types of supervised classifiers. A rule-based classifier uses the training set to construct rules as to how objects presented to it in future should be classified. A common way of formulating rules is as a decision-tree, where a rule is applied and depending on the outcome other rules branching off that rule are applied until a classification outcome is achieved.

### B. *Experimental Method*

To prepare training and testing of our classifier, we obtained datasets containing Skype, normal VoIP and non-VoIP traffic. The non-VoIP traffic comprised game traffic and traces from the University of Twente [11]. Once the datasets were obtained, features from the captured files were segmented into sliding windows of 1 to 10 seconds using 1 second increments and separated into two different data sets which are training set and testing set. After training the classifier with a training set of particular features, we then used the testing set to evaluate its performance.

In this research, we used Skype and non-VoIP captured files (in .pcap extension) from previous research [1]. Non-VoIP traffic was obtained from two 24 hour traces from the University of Twente [9] and three hours of game traffic from Swinburne University of Technology. Skype traffic comprised 6.8 hours obtained from 18 calls made across a multi-hop public network. A total of 58 MB of Skype traffic was collected for analysis, comprising approximately710000 packets. The trace from University of Twente comprised approximately 503 Mbytesand approximately 752000 flows.

As part of the project, normal VoIP traffic was captured from 15 calls that made up approximately 5 hours of communication between two 7960 series CISCO IP phones or a CISCO phone with home phone in normal PSTN network. Tcpdump which was installed in a bridge under FreeBSD and was used to capture normal VoIP traffic [7].

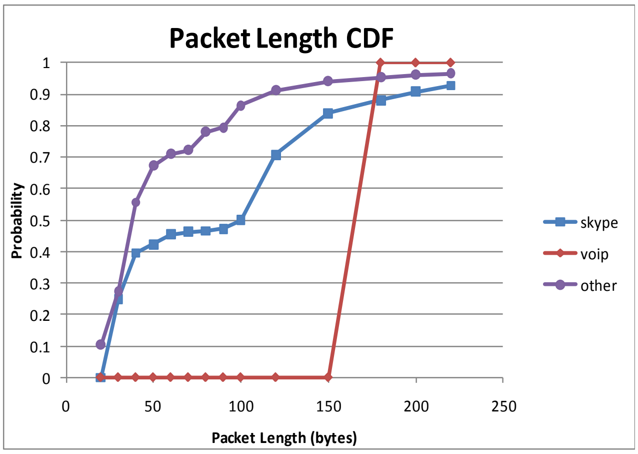

Figure 1. Packet Length Cumulative Distribution Function

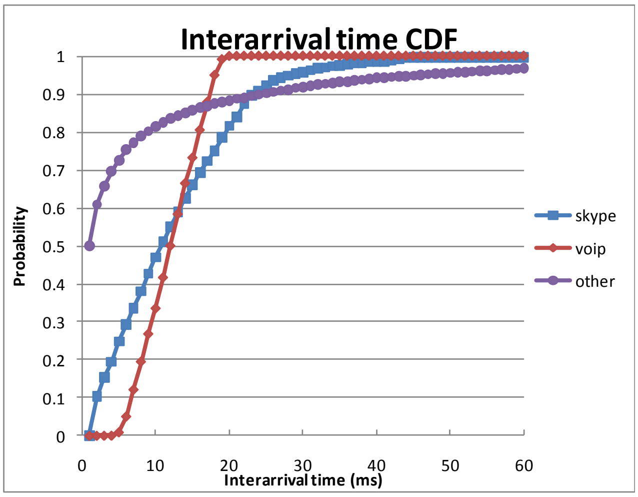

Figure 2. Packet Length Cumulative Distribution Function

All raw captured data was processed by Netmate to generate traffic flows in term of protocol, source IP address, destination IP address and port number. Each flow was segmented into subflows of 1 to 10 seconds. Those flows were then passed through WEKA to train the classifier.

We separated our data source into two different sets. The first half of all traffic was used to build up the training set and the other half was used for testing purposes. Our research also showed that Decision Tree classification using the J48 classifier is the most successful classifier. We also used the C4.5 Decision Tree Algorithm to choose suitable features.

### C. *Experimental Results*

#### 1. *Packet Length*

The first packet characteristic that we used is packet length. Skype packet length varies while normal VoIP traffic has fixed length packets. Other traffic packets are highly variable.

Our VoIP phones used Codec G.711u with the sample period of 20 ms. Packet size for this type of codec is 160 bytes/sample. Including IP header, each packet has a size of 200 bytes. By using UDP packet length (fixed at 180) and standard deviation of 0, it is possible to detect VoIP from non-VoIP traffic accurately.

Skype uses a different Codec that generates variable length packets in the range of 20 bytes to 200 bytes plus.

Online game traffic has different characteristics to VoIP. Around 70% of game traffic has the length of 40 bytes.

We used the mean and standard deviation of packet length to construct a classifier using WEKA [7]. Our results are shown in Table 1.

### 2. *Inter-arrival Time*

Another useful attribute that we can consider for our classifier is the inter-arrival time between packets. From a previous study [1], we know that the Skype Client generates packets at intervals that are a multiple of 16 milliseconds. We observed intervals times of 16, 32, 48 and 64 milliseconds. VoIP delay its traffic to generate multiple samples to send in a single packet. This delay is the sum of sample time and time to transfer each packet. Game traffic should be sent as fast as possible.

As we can see from Figure 2, most of the inter-arrival times between 'Other' packets is less than 10 millisecond while Skype ranges from 0 to 35 milliseconds of delay. VoIP delay is a longer than game traffic which is from 8 to 15 milliseconds. Calculating mean and standard deviation of inter-arrival time in statistic flows will help us identify Skype and VoIP traffic.

Applying this feature to train a classifier, our result is shown in Table 2.

### 3. *Combination of both features*

We also constructed a classifier that combined both packet length and inter-arrival time statistics. Performance is shown in Table 3. We see excellent recall for 10 second windows.

## IV. Conclusion

Our work focused on two main attributes to classify VoIP and Skype traffic. These were packet length (mean and standard deviation) and interval arrival-time. We also combined both these two groups of attributes in order to increase the power of classification.

Packet length statistics worked well to identify VoIP since VoIP has fixed packet length. Skype can also be classified because of its well defined statistical characteristics. From our testing, inter-arrival time is also a good feature to use for classification. A classifier that uses both packet length and inter-arrival time is able to identify the flow very accurately.

## Acknowledgments

Hoang would like to thank Philip Branch and Jason But for their supervision, Grenville Armitage for his support and CAIA for the opportunity to undertake an internship.

Table 1. Recall for a classifier based on packet length statistics

| Type of traffic | 1 second | 10 seconds |
|---|---|---|
| Skype | 96% | 98% |
| VoIP | 100% | 100% |
| Non-VoIP (game, other) | 92% | 97% |

Table 2. Recall for a classifier based on inter-arrival time statistics

| Type of traffic | 1 second | 10 seconds |
|---|---|---|
| Skype | 89% | 97% |
| VoIP | 93% | 99% |
| Non-VoIP (game, other) | 92.5% | 99.1% |

Table 3. Recall for a classifier based on packet length and inter-arrival time statistics

| Type of traffic | 1 second | 10 seconds |
|---|---|---|
| Skype | 95% | 99% |
| VoIP | 99% | 100% |
| Non-VoIP (game, other) | 93% | 99.3% |

## References

[1] P. Branch, A. Heyde and G. Armitage "Rapid Identification of Skype Traffic," in *ACM NOSSDAV 2009*, Williamsburg, Virginia, USA, 3-5 June 2009

[2] I.H Witten and E. Frank "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[3] N. Nilsson, "Introduction to Machine Learning," accessed August 2009 [Online] http://robotics.stanford.edu/people/nilsson/mlbook.html,

[4] T. Mitchell, "Machine Learning," McGraw Hill 1997

[5] CAIA, "Network Traffic based Application Identification," accessed August 2009 [Online] http://caia.swin.edu.au/urp/dstc/netai/,

[6] "Netmate", accessed August 2009, [Online] www.ip-measurement.org

[7] University of Waikato, "Weka", accessed August 2009 [Online] www.cs.waikato.ac.nz/ml/weka/,

[8] Tcpdump [Online] accessed August 2009, www.tcpdump.org/

[9] Wireshark [Online] *www.wireshark.org*, accessed August 2009

[10] Federal Communications Commission (US Govt), "Communications Assistance for Law Enforcement Act" accessed August 2009 [Online], www.fcc.gov/calea/ ,

[11] University of Twente, "Traffic Measurement Data Repository" [Online] http://traces.simpleweb.org/ , accessed February 2009

[12] N. Williams, S. Zander, S., G. Armitage, "A Preliminary Performance Comparison of Five Machine Learning Algorithms for Practical IP Traffic Flow Classification", ACM SIGCOMM Computer Communication Review, vol. 36 no. 5 pp. 7-15, 2008

[13] T. T. T. Nguyen, "A novel approach for practical real-time, machine learning based IP traffic classification", PhD thesis, Swinburne University of Technology, 2009.

[14] T. T. T. Nguyen and G. Armitage, "A Survey of Techniques for Internet Traffic Classification using Machine Learning", IEEE Communications Surveys & Tutorials, vol. 10 no. 4 pp. 56-76, 2008

.