Analysis of Skype and its Detection

Th. Rupachandra Singh, Irengbam Tilokchan Singh, Tejmani Sinam

Abstract*—This paper gives a complete analysis of Skype Traffic. Based on the analysis of Skype Traffic, we proposed a heuristic based detection method which classified the Skype Signaling and Skype Media Traffic. We properly categorized the Skype Media traffic as audio or video conversation. In this paper, we also propose a novel approach to identify VoIP Network Traffic in the first few seconds of initial state of communication. The proposed classifier works with Machine Learning Techniques based on the statistical features. The experimental results show that the proposed method can achieve over 99% accuracy for all testing dataset.

***Index Terms*— Skype; Network Traffic Analysis; Traffic Classification; Machine Learning**

I. INTRODUCTION

In general, network traffic classification is a fundamental process to classify the network traffic and identify the corresponding applications in modern network security systems, network monitoring, QoS and traffic engineering. So, there's a lot of interest among research community, network operators and even Government agencies, in identifying these applications. Traditional method of traffic classification are done based on the application port mapping which are assigned by IANA (*Internet Assigned Numbers Authority*), protocol format analysis and payload based matching approach. But today, emerging applications uses ephemeral, dynamic and random ports and encrypted payloads for obfuscation. So, the traditional methods of traffic classification (*port based prediction* and *payload based deep inspection method*) [1]–[4] are no longer effective and efficient. Most researchers are diverting away from these old techniques of classification and are adopting the statistical based classification techniques. These statistical based classifications were carried out using Machine Learning (ML) [5]–[11] and significant gains have studied from the previous methods.

With emerging trends and technologies, users nowadays are shifting from traditional phone call to VoIP applications. These applications are mostly encrypted; some like Skype uses P2P architectures and have the capability to traverse any network conditions. Some of VoIP applications that we have considered in our study are *Skype*, *Gtalk*, *Asterisk* and *Google+ Hangouts.*

In this paper, we give a detail analysis of Skype VoIP application. The analyses are performed on the network traces generated by them. We deal with signal and media traffic, so that the corresponding application can be detected from the network traffic. In the statistical proposed method, media traffic flows of a particular application are gathered first. These flows are further split into *sub-flows* using sliding windows. The term *sub-flow* is defined as subsets of a flow that have the same 5 tuple (*src ip, dst ip, src port, dst port* and *protocol*) with time based windows size (1, 3, 4, 5, 7 and 10 seconds) and are obtained by sliding windows. These windows are overlapping. Let us consider how to to obtain sub-flow window size of *5* second; the *1st* window start from *0* second to *5* second; the *2nd* window start from *1* second to *6* second; and *kth* window start from *k-1* to *k+4* second; thus these window are sliding with *1* second and overlapping with *4* second. The reasons for considering sub-flows are:

- *early classification of VoIP media Traffic* and
- *Classifier works in real time or on- the-fly.*

The security, monitoring and management systems need the prior information, but not the post-mortem report.

From these *sub-flows* we extract the statistical flow features. These features consist of packet's counter, packet size, minimum packet size, maximum packet size, the first and the second order statistics over packet size, and packet inter-arrival (*minimum, maximum, average and standard deviation*) are obtained using overlapping sliding window.

In our study, we use four supervised Machine Learning classification algorithm viz., *Decision Tree (C4.5), Naïve Bayes, Bayesian Belief* and *SVM* [12]–[15]. First of all, we build different classifier model based on *sub-flow* statistic with varying window sizes.

The rest of this paper is organized as follows. Section II covers the related work with more emphasis on statistics based Internet traffic classification approaches. Section III outlines the data used and how they are collected. Section IV describes the detailed analysis of Skype Traffics. Section V gives the detection methodologies that are proposed and discuss the performance measures. Section VII concludes the paper with some final remarks and suggestions of possible future work.

II. RELATED WORKS

Traditionally, network traffic classification were carried out base on IANA assigned reserved port numbers. Newer applications such as those using p2p have used ephemeral ports to bypass traffic and thus port-based classifications are ineffective. This was confirmed by Karagiannis et.al.[16] by identifying P2P on handcrafted signatures. Haffner et.al.[17] automated the construction of application signatures on trained sets by employing supervised machine learning techniques. Jeffrey Erman et.al.[18] showed the performance of k-means clustering better than DBSCAN and EM clustering algorithms. Jeffrey Erman et.al.[19] proposed a semi-supervised learning utilizing k-means clusters. This study was based on statistics of the flow. The k-means clustering has the drawback of assigning the number of cluster and the number of

Revised Version Manuscript Received on August 31, 2016.

Dr. Th. Rupachandra Singh, Department of Computer Science, Manipur University, Imphal, India.

Irengbam Tilokchan Singh, Department of Computer Science, Manipur University, Imphal, India.

Tejmani Sinam, Department of Computer Science, Manipur University, Imphal, India.

cluster to be formed cannot be predicted. So, Yu Wang et.al.[20] employed X-means clustering to their work. Although X-means is basically equivalent to k-means, it does not require the assignment of the number of clusters in advance. Xiang Li et.al.[21] applied Support Vector Machine learning based on flow statistics to identify and classify network applications. Nowadays, researchers are more or less attracted towards statistical based approach as it does not involve packet payload data. Many researchers attempted to build statistical classifier models based on full flow feature, but these full flow features approaches exhibit slower computational performance. Now a day, instead of full flow feature researchers have started the use of sub-flow features [5], [22]–[24]. Thus the packet size and inter-arrival time are more effective measurable features in early classification of network flows [5], [25], [26].

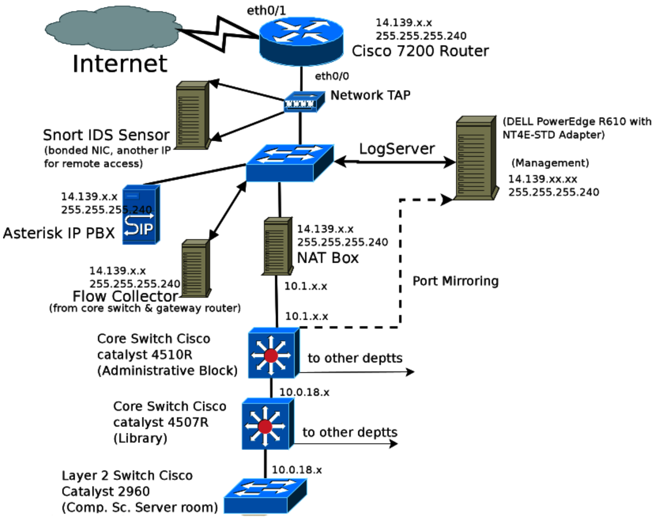

Fig.1. Data Collection Architecture

III. DATA COLLECTION

Network traces are collected from our test-bed, at the edge of our University Network (Figure 1) and from publicly available traces of *Tstat* [27] [28]. Figure 2 shows the test-bed is setup at Network Security Lab, M.U. (*Manipur University*) which generate various VoIP application traces. And using gt's [29] method we collect ground truth application traces. A Napatech data capture card, NT4E-STD [30] was used to capture traces on our log server at the edge of our University Network (*Figure 1*). An Asterisk based VoIP server is also running at the public domain to collect the Asterisk based VoIP applications traces. We also collected various types of Skype and non-Skype traces such as voice, video, silence call, call within LAN and WAN, etc.

Data were Data were generated using the VoIP clients such as Skype (Beta) version 2.2.0.35, linphone 3.5.2 (Windows 7), linphone 3.3.2 (linux mint 13), sipdroid 2.7 beta, Ekiga Softphone 3.3.2, 3CXPhone 6.0.26523.0 (Windows 7), Gtalk in Google Chrome v20.0.1132.57, Gtalk in Google Chrome v23.0.1271.91, Gtalk with Empathy 3.4.2.3, Google+ Hangout and Asterisk 11.0.0 beta1 using Android 4.0.3 (ICS), Android 4.1-4.3 (Jelly Bean), Windows 7, Linux-mint 13 and Ubuntu 12.04-14.04.

For the experiment, we were able to collect ≈ 32 GB of VoIP traces including Skype's 14.71 GB, spread over 3-4 months during 2011- 2013. And another 3.8 GB of Skype's anonymized traces was obtained from Tstat. We only downloaded end to end Skype UDP traffic from Tstat. From these anonymized traces we extracted about ≈ 586 GB of packet size as derived from the header IP lengths.

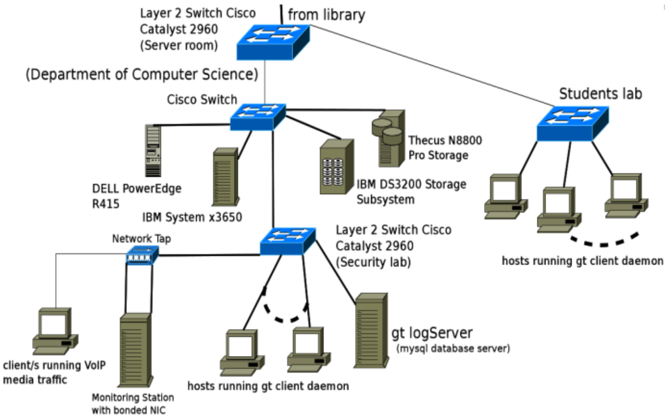
Layer 2 Switch Cisco Catalyst 2960 (Server room)
from library
(Department of Computer Science)
Cisco Switch
Students lab
Thecus N8800 Pro Storage
DELL PowerEdge R415
IBM DS3200 Storage Subsystem
IBM System x3650
Network Tap
Layer 2 Switch Cisco Catalyst 2960 (Security lab)
hosts running gt client daemon
gt logServer
(mysql database server)
client/s running VoIP media traffic
Monitoring Station with bonded NIC
hosts running gt client daemon

Fig. 2. Test-bed Setup at Network Security LAB

IV. ANALYSIS OF SKYPE

Skype is one of the most popular VoIP (*Voice over Internet Protocol*) applications. Basically, VoIP is the transmission of message (*voice* or *video*) traffic over IP-based networks. All conversations are sent as data packets over the network. VoIP traffic usually consists of signalling and media. The media protocols are used to transmit media such as audio and video over IP networks. Signalling protocols are responsible for the establishment, preservation and tearing down of call sessions. They are also responsible for the negotiation of session parameters such as codecs, tones, bandwidth capabilities, etc. Skype is propriety p2p VoIP network and is designed to work out of the box on modern networks, and has no problems working behind a NAT device or other firewalls. Skype is based on proprietary protocols, which make extensive use of cryptography, obfuscation, and anti-reverse engineering procedures making casual eavesdropping or impersonation all but impossible. The architecture of Skype is still in the black box, but the Network Administrator/Analysis/Reverse Engineers have said that it is an overlay peer-to-peer network. The peer-to-peer Skype network consist of SC (*Skype Client: main application runner*), SN (*Super Node: act as Relay Node Distributed*) and LS (*Login Server: a centralized Skype Server*). In this section, we analysed the behaviour characteristic of Skype from the network traces generated under different circumstances.

A. Skype Signalling

Skype used different type of signal mechanism. *Skype Client* (SC) exchanges packets with *Skype Super-Node* (SN) and the other *Skype Client*. All the activities of a *Skype Client (SC)* can be seen /found on the network traces. A Skype client application may *start-up, login, online, active, idle, call, IM, etc*. The activities of a Skype application are sent to the SNs and informed to other SCs. So some signal are for *status information, some are for keep-alive status to traverse NAT, some are for initial call set-up, some are for call hang-up etc*. There are two type of Skype signal flow base on connectivity time pattern, i.e. *Short and long packets exchanges with SNs*. Base on connectivity behaviour pattern, it have been classified in to three types, viz

i. *burst-connection*: bunch of 25 instantaneous connection in the first few second of application start-up,

ii. *activities status report:* maximum of 5 new connection with SNs for every status report of online, offline, idle activities and

iii. *during conversation*: two periodical connection with SNs during voice or video call conversation between clients.

Signal Flow characteristic: All the Skype Signal flows have a deterministic property. The last four bits of the 3rd bytes of the UDP packets payload have found with magic value (2, 3, 5, 7, 11, 13 and 15). The type 13 is found only on the media flows. Based on this property, we proposed a heuristic to identify a flow as Skype signal traffic or Skype media traffic.

Skype Client (SC) to Skype Client (SC) TCP Signal characteristic: During a Skype-to-Skype call, Skype Client (SC) usually transfer the media in UDP protocol and in between the call, SCs have TCP connection with the same *src-dst-port* as used in the UDP protocol. If we identify the UDP flow as Skype than we can conclude that the TCP flow having the same *src-port* endpoint as Skype.

B. *Skype Media Traffic*

The major volumes of Skype traffic are the media traffic. If we are able to know the media flows of Skype, than the majority portion are classified. By default the Skype media traffic are usually UDP. It conversed to TCP upon blocking UDP. As we have mention earlier, the four bits of the 3rd bytes of UDP payload for media traffic are reserve as 13. Thus, the prediction of Skype UDP media traffic is quite easier, compare with TCP. Based on packets size distribution, we can broadly classified Skype UDP media flow as *voice call* or *video call*. Usually *voice call* packets are small with compare with *video call* records.

There are two type of media flow traffic in Skype based on connection pattern, viz *direct call conversation* and *relay call conversation*. The *direct call conversation* are set up when a SC able to communicate with other SC application port map. If the client cannot communicate with the Skype application reserve port, then the call are relayed through a relay server. This type of call are said to be *relay call conversation*.

C. Analysis of Skype Traffic with various codec

We generated various traces of Skype application. The experimentation was carried out with various codecs used by the VoIP Skype application. List of 11 codec that are supported by Skype application are SILK_V3, SILK_WB_V3, SILK_MB_V3, SILK_NB_V3, DECT, SVOPC, SOVPC_SB, NWC, G729, PCMA and PCMU. We generate various Skype VoIP traffic be selecting each of the codec one by one. We analysis the Skype VoIP traffic with inter-packets-arrival time, packet generation rate, packet size distribution or histogram of packet size and trying to correlate the signal information traffic.

Table I. Skype Codec Property

Sl. No.	**CODEC**	**IN-BOUND or OUT-BOUND**			**Inter-packets-arival-time (mn,max)**
		pps	***KBps***	***(min, max, most frequent pkts size)***	
1	PCMA	50	9	31, 614, 197	0.0, 1.003
2	PCMU	50	10	31, 715, 207	0.0, 8.169
3	G729	50	2-3	31, 631, 57	9.5e-7, 1.53
4	SVOPC	60/20/10	9-10	31, 723, 45	0.0, 9.10
5	SVOPC-SB	20-60	1-8	31, 1029.45	0.0, 8.942
6	SILK-V3	20-55	2-6	31,1035, 95/98	0.0, 12.232
7	SILK-WB-V3	55	4-6	31,1038, 88/89/90	0.0, 10.049
8	SILK-MB-V3	25-55	1-4	31, 1036, 67	0.0,10.023
9	SILK-NB-V3	55/20	1-3	31, 559/1038, 61/60.62	0.0, 14.894
10	NWC	50/20/18	10/2/1	31, 670, 197	0.0, 9.693
11	DECT	18-50	9-15	31, 1015. 197	0.0, 10.004

We observed that one of the *deterministic properties* of Skype VoIP traffic is the minimum packet size is 31bytes. The 31 bytes packet size found correct for the entire UDP stream. The most frequent packets sizes are vary with different codec of an Application. Even though different application used same codec, there is a deterministic property like most frequent packet size, minimum and maximum packets size, this is due to the fact that different application used different packetization property. For Skype the minimum packet size for UDP and TCP are less than 80 bytes. For other VoIP like Gtalk the minimum packet size is greater than 80 bytes.

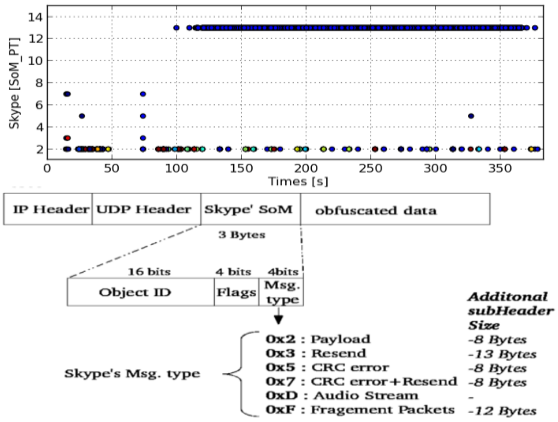

Fig. 3. Skype UDP Packets Structure

D. *Analysis of Skype behaviour under various conditions, viz blocking certain ports, blocking UDP protocol, blocking DNS protocol, etc.*

We collected different Skype call traces under different network scenarios. Some session are restricted to use only TCP, some session with only UDP and some session by blocking the application port number. The protocol filtering is tested at the end-Point machine as well as the NAT-BOX firewall. For the End-Point machine, the rule is added at the INPUT and OUTPUT chain rule, but for the firewall at the NAT-BOX the rule is implemented at the FORWARD chain rule.

The lists of experiment that have performed at our test bed are as under:

- Block the DNS protocol, and test whether the Skype can able to join the Skype overlay network and can start a conversation with the other Skype Client peers.
- If we drop all the UDP protocol except DNS, the Skype Client can still login to the Skype Network.
- Even if we drop all the UDP packets including the DNS protocol, it can still communicate with other Skype clients. In this case Skype used with TCP protocol for transport protocol.
- Block the UDP port number used by the Skype Client, can still communicates with other peers.
- If we block the DNS and UDP port number used by the Skype application, it can still connect and communicate with other peers.
- If we drop all the DNS and port number greater than 1024 for both TCP and UDP, at last it cannot connect to the Skype network.

There are two type of TCP connection in the Skype communication.

i. *DIRECT_TCP*: The two Skype Client have direct TCP connection.

ii. *RELAY_TCP*: The two Skype clients have no direct TCP connection but connected through a Relay Server.

A (Skype Client) make a call to B (Skype Client), by blocking UDP protocol in A. A DIRECT_TCP connection is SYN/started/initiated by B (S_C) to A (S_C). The port number used by the (A and B) clients are same as the port that have used by the Skype Application' UDP protocol. If B (S_C) is unable to set up direct connection with A (S_C), than the A (S_C) and B (S_C) are communicated through a relay server C (R_S). Both A (S_C) and B (S_C) have TCP communication with C (R_S), in this case the A (S_C) and B (S_C) have RELAY_TCP connection. The relay communication was done only when A (S_C) and B (S_C) cannot select their application port numbers. Usually the minimum packet sizes of TCP packets are 52 bytes.

V. DETECTION METHODOLOGY

A. *Heuristic based Approaches*

Heuristic based approaches for detection of Skype are purely based on the Skype packets structure and its distribution in communication (figure 3). And based on packets size we differentiate the voice and video traffic. We proposed two type of heuristic: *(a.) Signal and media detection heuristic* and *(b.) voice and video detection heuristic*.

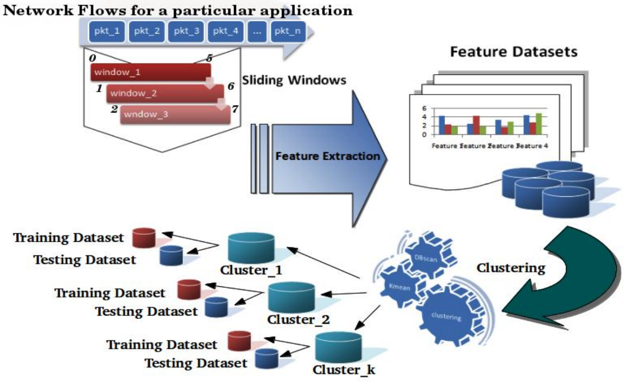

Fig.4. Machine Learning: Application Feature Datasets extraction to gather training and testing datasets, proportionality sampling by using clustering

B. *Statistic based Approaches using Machine Learning*

Several significant studies have previously been carried out on traffic classification based on Machine Learning (ML) [5]–[11]. Some are focused on clustering techniques which are unsupervised Machine Learning algorithms and some are based on supervised Machine Learning method, which deals with training the classifier with known datasets. The method proposed in this paper is a hybrid approach based on the combination of both unsupervised and supervised methods.

The traces used in our study were captured on the client side and at the edge of our University Network during 2011-2013. The ground truth traces are categorized into two set (*training* and *testing*). From these two set, we extracted the feature datasets for each application. And *k*-mean clustering was performed on each application datasets to group the training set for each application. The value of *k* is determined from the result of DBSCAN clustering. Lastly, we balance training and testing set by acquiring the data proportionality of each cluster of each application.

In our study, we use four supervised Machine Learning classification algorithm viz., *Decision Tree (C4.5)*, *Naive Bayes*, *Bayesian Belief* and *SVM* [12]–[15]. First of all, we build different classifier model based on *sub-flow* statistic of 22 attributes feature derived by varying window size of 1, 3, 5, 7 and 10 seconds. We analyze the classifier models with deferring windows. We also apply model to test Tstat trace [27], [28].

We extracted the statistical *sub-flow* features of the media traffic using overlapping sliding windows as shown in Figure 4. From the training datasets of Figure 4 build the supervised Machine Learning Classifier model. Figures 5 show the proposed machine learning traffic classification model building and the classifier which classify the offline datasets. Our system goes through the following stages:-

1) *Training and testing datasets preparation,*
2) *Model building* and Classification.

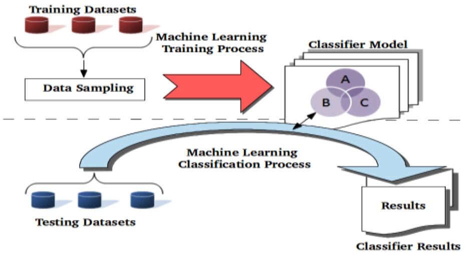

Fig.5. Training the Machine Learning traffic classifier and classification of offline datasets

TABLE II SAMPLE DATASETS

Model	Data sample points
Skype vs (Gtalk+Hangouts+Asterisk)	6000 (Skype) + 6000 (Other)
Asterisk vs (Gtalk+Hangouts+Skype)	6000 (Asterisk) + 6000 (Other)
Gtalk/Hangout vs (Skype+ Asterisk)	6000 (Gtalk/Hangout)+6000(Other)

For each of the application, a sample of 6,000 tuples is randomly selected from the original data. So, altogether we use 18,000 tuples which is used to build the model. The training dataset is shown in table II. Similarly, testing data points are selected from the separate testing trace. So, altogether the testing data points consist of 18,000 data points contributed by the three applications.

TABLE III. Comparison of Performance Measurement with Different ML Algorithms (Multiclass Classifiers)

Window	C4.5	BBN	NB	SVM
1-second	99.18%	95.99%	42.23%	98.25%
3-second	**99.47%**	96.61%	37.88%	99.05%
5-second	99.30%	96.64%	40.81%	97.63%
7-second	99.24%	97.32%	42.88%	96.56%
10-second	99.36%	98.06%	48.06%	98.93%

We build various classifier based on the following machine learning algorithms: *NB (Naive Bayes), BBN (Bayesian Belief Network), C4.5* and *SVM (Support Vector Machine)*. Implementations of all these machine learning algorithms are done using *WEKA*[31]. The *sub-flow* performances are compared with different sliding windows size. Table III shows result of testing the classifier model with window size of 1, 3, 5, 7 and 10 second.

From table III, we can see that C4.5 classifier is the best among the ML classifiers used in our experiments. C4.5 classifier model on 3-second sliding window achieved the highest result of 99.47%. So, C4.5 classifier algorithm is determined as the best classifier model with 3-seconds window and chosen for further studies. The precision and recall value for 3-seconds window based on C4.5 classifier is shown in the table IV.

TABLE IV. Precision And Recall Value Calculated Based on C4.5 Classifier Model With 3 Second Window Size

Class	Precision	Recall
Astrisk	0.99	0.99
Skype	0.99	0.99
Gtalk	0.99	0.99

VI. CONCLUSION AND FUTURE WORKS

We design a network traffic classifier based on the statistical features extracted from network flows. Instead of deriving the statistical characteristics per flow, our models make use of features extracted from the first few seconds of each flows. The first few seconds of each flow is divided into overlapping time-based windows. This approach enables our classifier to classify each flow early. Our approach of network traffic classification utilize the statistical information contained in the packet header and do not utilize any information contained in the payload. Classification involving payload is not always feasible due to encryption and also due security and privacy concerns.

We give a comparative analysis of the result on the said approach based on the classification algorithms (Decision tree (C4.5), Naive Bayes, Bayesian Belief Network and SVM). The experimental results show that the proposed method can achieve over 99% accuracy for all testing dataset.

Based on our experience of building a number of off-line classifiers we have developed a number of online network traffic classifiers based on Machine Learning and heuristics techniques. These classifiers will be deployed in our university network and the experience gained will be utilized to further refine our classifiers. Our group is in the process of obtaining more datasets from various sources which we intend to use to test our classifier for robustness.

ACKNOWLEDGMENT

This work was supported by a grant from the Department of Electronics and Information Technology (DeitY), Ministry of Communication and Information Technology, Government of India by funding the project “Network Traffic Classification and Analysis”.

REFERENCES

1. S. Sen, O. Spatscheck, and D. Wang, “Accurate, scalable in- network identification of p2p traffic using application signatures,” in Proceedings of the 13th International Conference on World Wide Web. New York, NY, USA: ACM, 2004, pp. 512–521.
2. “l7-filter application layer packet classifier for linux,” 2009, http: //l7filter.sourceforge.net.
3. T. Sinam, I. T. Singh, P. Lamabam, and N. N. Devi, “An efficient technique for detecting skype flows in udp media streams,” in Advanced Networks and Telecommuncations Systems (ANTS), 2013 IEEE International Conference, Dec 2013, pp. 1–6.
4. T. Sinam, I. T. Singh, P. Lamabam, N. N. Devi, and S. Nandi, “A technique for classification of voip flows in udp media streams using voip signalling traffic,” in Advance Computing Conference (IACC), 2014 IEEE International, Feb 2014, pp. 354–359.
5. T. Sinam, N. N. Devi, P. Lamabam, I. T. Singh and S. Nandi, “Early Detection of VoIP Network Flows based on Sub-Flow Statistical Characteristics of Flows using Machine Learning Techniques,” in Advanced Networks and Telecommuncations Systems (ANTS), 2014 IEEE International Conference, Dec 2014.
6. L. Grimaudo, M. Mellia, E. Baralis, and R. Keralapura, “Select: Self-learning classifier for internet traffic,” IEEE Transactions on Network and Service Management, vol. 11, no. 2, pp. 144–157, 2014.
7. T. T. Nguyen and G. Armitage, “A survey of techniques for internet traffic classification using machine learning,” Commun. Surveys Tuts., vol. 10, no. 4, pp. 56–76, Oct. 2008.
8. J. Chandrakant and D. Lokhande Shashikant, “Analysis of early traffic processing and comparison of machine learning algorithms for real time internet traffic identification using statistical approach,” in Advanced Computing, Networking and Informatics- Volume 2, ser. Smart Innovation, Systems and Technologies, M. Kumar Kundu, D. P. Mohapatra, A. Konar, and A. Chakraborty, Eds. Springer International Publishing, 2014, vol. 28, pp. 577–587.
9. R. Yan and R. Liu, “Principal component analysis based network traffic classification,” JCP, vol. 9, no. 5, pp. 1234–1240, 2014.
10. J. M. Reddy and C. Hota, “P2p traffic classification using ensemble learning,” in Proceedings of the 5th IBM Collaborative Academia Research Exchange Workshop, ser. I-CARE ’13. New York, NY, USA: ACM, 2013, pp. 14:1–14:4.
11. M. Korczynski and A. Duda, “Markov chain fingerprinting to classify encrypted traffic,” in IEEE Conference on

Computer Communikations, INFOCOM , Toronto, Canada, April 27 - May 2, 2014. IEEE, 2014, pp. 781–789.

12. "libsvm-3.0," http://www.csie.ntu.edu.tw/~ cjli n/libsvm/.
13. N. Cristianini and J. Shawe-Taylor, An Introduction to support Vector Machines and other Kernel-based Learning Methods. Cambridge University Press, 2003.
14. I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and techniques. Elsevier Inc., 2005.
15. J. Han and M. Kamber, Data Mining: Concepts and Techniques. Elsevier Inc., 2006.
16. T. Karagiannis, A. Broido, N. Brownlee, K. C. Claffy, and M. Faloutsos, "Is p2p dying or just hiding?" in Proceedings of the GLOBECOM 2004 Conference. IEEE Computer Society Press, November 2004.
17. P. Haffner, S. Sen, O. Spatscheck, and D. Wang, "Acas: Automated construction of application signatures," in Proceedings of the 2005 ACM SIGCOMM Workshop on Mining Network Data, ser. MineNet '05. New York, NY, USA: ACM, 2005, pp. 197–202.
18. J. Erman, A. Mahanti, M. F. Arlitt, I. Cohen, and C. L. Williamson, "Semi-supervised network traffic classification," in SIGMETRICS, 2007, pp. 369–370.
19. J. Erman, M. Arlitt, and A. Mahanti, "Traffic classification using clustering algorithms," in Proceedings of the 2006 SIGCOMM Workshop on Mining Network Data, ser. MineNet '06. New York, NY, USA: ACM, 2006, pp. 281–286.
20. Y. Wang, Y. Xiang, and S.-Z. Yu, "An automatic application signature construction system for unknown traffic." Concurrency and Computation: Practice and Experience, vol. 22, no. 13, pp. 1927–1944.
21. X. Li, F. Qi, D. Xu, and X. Qiu, "An internet traffic classification method based on semi-supervised support vector machine." in ICC. IEEE, 2011, pp. 1–50.
22. T. N. Thuy T. and G. Armitage, "Training on multiple sub-flows to optimise the use of machine learning classifiers in real-world ip networks," in in Proceedings of the IEEE 31st Conference on Local Computer Networks, 2006.
23. S. Zander, T. T. T. Nguyen, and G. J. Armitage, "Sub-flow packet sampling for scalable ml classification of interactive traffic," in LCN, 37th Annual IEEE Conference on Local Computer Networks. Clearwater Beach, FL, USA: IEEE, October 22-25 2012, pp. 68–75.
24. G. Xie, M. Iliofotou, R. Keralapura, M. Faloutsos, and A. Nucci, "Sub-flow: Towards practical flow-level traffic classification," in Proceedings of the IEEE INFOCOM. Orlando, FL, USA: IEEE, March 25-30 2012, pp. 2541–2545.
25. A. Este, F. Gringoli, and L. Salgarelli, "On the stability of the information carried by traffic flow features at the packet level," SIGCOMM Comput. Commun. Rev., vol. 39, no. 3, pp. 13–18, Jun. 2009.
26. L. Peng, H. Zhang, B. Yang, and Y. Chen, "Feature evaluation for early stage internet traffic identification," in Algorithms and Architectures for Parallel Processing, ser. Lecture Notes in Computer Science, X.-h. Sun, W. Qu, I. Stojmenovic, W. Zhou, Z. Li, H. Guo, G. Min, T. Yang, Y. Wu, and L. Liu, Eds. Springer International Publishing, 2014, vol. 8630, pp. 511–525.
27. "Tstat - skype traces," http://tstat.tlc.polito.it/ traces-skype.shtml.
28. "Tstat - tcp statistic and analysis tool," http://tstat.tlc.polito.it/index.shtml.
29. F. Gringoli, L. Salgarelli, M. Dusi, N. Cascarano, F. Risso, and kc Claffy, "Gt: picking up the truth from the ground for internet traffic," Computer Communication Review, vol. 39, no. 5, pp. 12–18, 2009.
30. "Napatech," http://www.napatech.com/.
31. "Weka3.6.2," 2011, http://www.cs.waikato.ac. nz /ml/weka.
32. www.halcyon.com/pub/journals/21ps03-vidmar

AUTHOR PROFILE

Dr. Thounaojam Rupachandra Singh, Assistant Professor, Department of Computer Science, Manipur University**,** MCA**,** PhD (IT)**,** UGC-NET**.** Paper published:

i. **Th. Rupachandra Singh**, , Kh. Manglem Singh, Sudipta Roy, "Video watermarking scheme based on visual cryptography and scene change detection**"** International Journal of electronics and Communications (AEU), 8/2013 Volume 67, (AEU) 67(2013)645-651.

ii. **Th. Rupachandra Singh**, , Kh. Manglem Singh, Sudipta Roy, **"**Robust Video Watermarking Scheme Based on Visual Cryptography", 2012 World Congress on Information and Communication Technologies, IEEE Xplore, 978-1-4673-4804-1©2012IEEE, pages 873-888.

iii. **Th. Rupachandra Singh**, , Kh. Manglem Singh, Sudipta Roy, "Image Watermarking Scheme based on Visual Cryptography in Discrete Wavelet Transform**"** International Journal of Computer Applications (0975 – 8887) Volume 39 – No. 1, February 2012.

iv. Kh. Manglem Singh, **Th. Rupachandra Singh**, O. Imocha Singh and T. Romen Singh " A Blind Video Watermarking Scheme based on Scene Change Detection" Indo-US Conference & Workshop on CYBER SECURITY, CYBER CRIME & CYBER FORENSICS, August 19-21, 2009, Kochi, India.

v. T. Romen Singh, O.Imocha Singh , Kh. Manglem Singh , Tejmani Sinam and **Th. Rupachandra Singh** "Image Enhancement by Adaptive Power-Law Transformations**"** Bahria University Journal of Information and Communication Technology Volume 3 Issue 1 (BUJICT 2010), ISSN 1999-4974.

vi. T.Romen Singh, O.Imocha Singh , Kh. Manglem Singh , Tejmani Sinam and **Th. Rupachandra Singh "**Image Magnification based on directed linear interpolation**",** International Journal of Computer Sciences and Engineering Systems(IJCSES) China Vol. 3, No. 4. October, 2009, CSES International © 2009 ISSN 0973-4406.

vii. T.Romen Singh, O.Imocha Singh , Kh. Manglem Singh , Tejmani Sinam and **Th. Rupachandra Singh "**Image Magnification based on directed linear interpolation**" NCC 2009, I**IT Guwahati.

Irengbam Tilokchan Singh, Department of Computer Science, Manipur University, Imphal, India, MCA, UGC-NET. Paper published:

i. T. Sinam, **I. T. Singh**, P. Lamabam, and N. N. Devi, "An efficient technique for detecting skype flows in udp media streams," in Advanced Networks and Telecommuncations Systems (ANTS), 2013 IEEE International Conference, Dec 2013, pp. 1–6.

ii. T. Sinam, **I. T. Singh**, P. Lamabam, N. N. Devi, and S. Nandi, "A technique for classification of voip flows in UDP media streams using voip signalling traffic," in Advance Computing Conference (IACC), 2014 IEEE International, Feb 2014, pp. 354–359.

iii. T. Sinam, N. N. Devi, P. Lamabam, **I. T. Singh** and S. Nandi, "Early Detection of VoIP Network Flows based on Sub-Flow Statistical Characteristics of Flows using Machine Learning Techniques," in Advanced Networks and Telecommuncations Systems (ANTS), 2014 IEEE International Conference, Dec 2014.

Tejmani Sinam, HOD**,** Associate Professor, Department of Computer Science, Manipur University, and Paper published:

i. OI Singh, **T Sinam**, O James, TR Singh, "Local Contrast and Mean Thresholding in Image Binarization", International Journal of Computer Applications 51 (6)

ii. T. Romen Singh, O.Imocha Singh , Kh. Manglem Singh , **Tejmani Sinam** and Th. Rupachandra Singh "Image Enhancement by Adaptive Power-Law Transformations**"** Bahria University Journal of Information and Communication Technology Volume 3 Issue 1 (BUJICT 2010), ISSN 1999-4974

iii. TR Singh, OI Singh, KM Singh, **T Sinam**, TR Singh, "Image Magnification based on directed linear interpolation". International Journal of Computer Sciences and Engineering Systems(IJCSES)

iv. **T. Sinam**, I. T. Singh, P. Lamabam, and N. N. Devi, "An efficient technique for detecting skype flows in udp media streams," in Advanced Networks and Telecommuncations Systems (ANTS), 2013 IEEE International Conference, Dec 2013, pp. 1–6.

v. **T. Sinam**, I. T. Singh, P. Lamabam, N. N. Devi, and S. Nandi, "A technique for classification of voip flows in udp media streams using voip signalling traffic," in Advance Computing Conference (IACC), 2014 IEEE International, Feb 2014, pp. 354–359.

vi. **T. Sinam**, N. N. Devi, P. Lamabam, I. T. Singh and S. Nandi, "Early Detection of VoIP Network Flows based on Sub-Flow Statistical Characteristics of Flows using Machine Learning Techniques," in Advanced Networks and Telecommuncations Systems (ANTS), 2014 IEEE International Conference, Dec 2014.