Lightweight Internet Traffic Classification based on Packet Level Hidden Markov Models

Naveed Akhtar
Department of Electrical Engineering University of Engineering and Technology Lahore, Pakistan

Dr. Muhammad Kamran
Department of Electrical Engineering University of Engineering and Technology Lahore, Pakistan

Abstract—**During the last decade, Internet traffic classification finds its importance not only to safeguard the integrity and security of network resources, but also to ensure the quality of service for business critical applications by optimizing existing network resources. But optimization at first place requires correct identification of different traffic flows. In this paper, we have suggested a framework based on Hidden Markov Model, which will use Internet Packet intrinsic statistical characteristics for traffic classification. The packet inspection based on statistical analysis of its different characteristics has helped to reduce overall computational complexity. Generally, the major challenges associated with any internet traffic classifier are: 1) the limitation to accurately identify encrypted traffic when classification is performed using traditional port based techniques; 2) overall computational complexity, and 3) to achieve high accuracy in traffic identification. Our methodology takes advantage of internet packet statistical characteristics in terms of its size and their inter arrival time in order to model different traffic flows. For experimental results, the data set of mostly used internet applications was used. The proposed HMM models best fit the observed traffic with high accuracy. Achieved traffic identification accuracy was 91% for packet size classifier whereas it was 82% for inter packet time based classifier.**

Keywords—Hidden Markov model; traffic classification; network security; deep packet inspection; internet traffic modeling; Internet of Things

I. Introduction

Rapid developments in multimedia and broadband applications have made traffic classification a difficult subject, but over the years it has drawn significant importance [1]-[5] among researchers. Use of non-standard ports, user privacy and huge traffic load on the network is creating major bottlenecks to some of the developed techniques. Traditional port based classification techniques are not reliable and cannot identify encrypted traffic. Statistical analysis based deep packet inspection approaches have proven to be more robust and efficient to handle encrypted traffic, which have made it a fertile research area.

Network traffic classification is fundamental to number of network activities, including its management, security, planning and quality of service provisioning [6]. The prerequisite for Internet traffic classification is packet inspection. However, strict privacy policies and heavy network load coupled with high processing and infrastructure requirements for deep packet inspection engines have made it difficult to implement. Statistical analysis based packet inspection approaches have been very effective for encryption and protocol obfuscation. But still real time traffic classification and complexity of existing solutions is a big challenge. The debate for optimal technique for traffic classification is still open and with the emergence of new multimedia broadband applications, like Peer to Peer, Internet Protocol based Television and online Games, it has become very difficult for traditional classifiers to identify different traffic flows [7], [8]. The researchers have responded to this difficulty by working out different methods of internet traffic classification based on application level usage patterns and customer behavior. The authors [9], [10] have modeled internet traffic by using a stochastic process in which internet traffic has a self-similar character in nature. The overall behavior for this model was observed for different traffic flows in various network architectures.

II. Related Work

In port based identification techniques, each application is having a unique port number at the server side, and various applications are detected by doing analysis of TCP and UDP [27] traffic. But before applying any traffic engineering rules, the captured port numbers are compared against their default ports [28] in order to validate correct port identification. But the rapid advancements in various applications, some authors have assigned port numbers other than their default port numbers. Bit-torrent [29] is one of such applications which use different port numbers. Due to such cases, the port based detection could not identify 30% of Internet traffic [30]-[32]. Table 1 below shows some examples of different ports assigned to different application by Internet Address Assignment Network Authority.

TABLE I. IANA Assigned Ports to Various Applications

Assigned Port	Application
20	FTP Data
21	FTP Control
22	SSH
23	Telnet
53	DNS
80	HTTP
110	POP3
123	NTP
161	SNMP

Nguyen and Armitage [11], covers the detailed and comprehensive work about traffic classification up to 2008. But due to the failure of two main packet classification techniques: 1) mapping of transport layer source and destination ports; and 2) payload signature based recognition, the researchers have focused their work on traffic classification using statistical and Machine Learning techniques. Nguyen, Thuy TT and Grenville Armitage [11] used machine-learning technique to analyze interactive IP traffic. W. Li and A. W. Moore [12] have suggested machine learning approach based on Naive Bayes and C4.5 decision tree algorithms, which accurately classify internet traffic by collecting different features at the start of internet traffic flows. There are number of packet scanning applications which are implemented across different networks, and they are capable of doing packet inspection, like SNORT [24], [25] and Linux L-7 (Layer-7) filter. One very important key area is Network security, where the intrusion takes place to take over system resources and causing denial of service for end users. To mitigate such attacks, authors [26] have suggested passing over the entire traffic through a firewall where all rules have been defined. The implementations [14], [15] works on statistical properties of different flows, i.e. IPT (Inter packet time) and PS (Packet size). Similarly, HMM implementation [16] covers the comparative analysis of different HMMs with other techniques of traffic classification. The researchers also applied other statistical methods [17], [18] to address the problem of traffic classification in IP networks.

III. Hidden Markov Model

These are stateful statistical models which are based on statistical principles of Markov Chain, which is a stochastic process where one state depends on the other state and are linked with each other through state transition probabilities. HMM can be represented at a high level by following variables:

1) The hidden variables with their temporal evolution follow a Markov chain, i.e. $x_n = s_1, s_2 \ldots, s_N$ represents the (hidden) state at discrete time n with N representing the number of states.

2) The observable variables which stochastically depends on the hidden state, i.e. $y_n = O_1, O_2 \ldots, O_M$, it represents the observable variables at discrete time n with M being the number of observable variables.

$\lambda = (\mu, A, B)$ represents key characteristics of Hidden Markov Model, Where

1) μ is the initial state distribution, i.e. $\mu_i = P(x_i = s_i)$

2) A is $N \times N$ transition Matrix, where N is representing number of states 1,2…,N.

3) B is $N \times M$ observable generation Matrix, where M is the observation matrix and it could be discrete or continuous in nature. Each observation can be described by different distributions and all these distributions are log-concave in nature.

The probability of being in any specific state while considering the same Markov Chain λ at a certain time t is as under

$$P(s_t = i \mid s_{t-1} = j, s_{t-2} = k, ..., \lambda) = P(s_t = i \mid s_{t-1} = j, \lambda) \quad (1)$$

HMM based estimation model was developed by using HMM estimation capabilities (learning, modeling, and prediction) [19], both for PS and IPT separately. The traffic classification model [9] recognizes the distinct behavior patterns of various flows. In HMM implementation [13], first few packets are used to train the model and to classify each flow at an early stage. The basic HMM structure learns the characteristics of initial packets of different flows and afterwards, the statistical properties of the complete sequence are determined by observing packet size and inter packet time. Following four mostly used application classes:

1) Live streaming (YouTube)

2) Email services

3) Online game

4) Voice services (Skype) were used to develop the model.

IV. Methodology and Approach

In order to Model different traffic flows, we focused four mostly used applications. These applications were represented by four different states based on their statistical properties, i.e. packet size and inter packet time. These applications were selected based on their usage and complexity. The classifier block diagram is as under.

The related traffic was generated from dedicated network machines and it was captured on a server placed in Network Operation Center. The considered traffic statistical parameters, i.e. packet size and inter packet mean and standard deviation was calculated using MATLAB and is shown in Table 2.

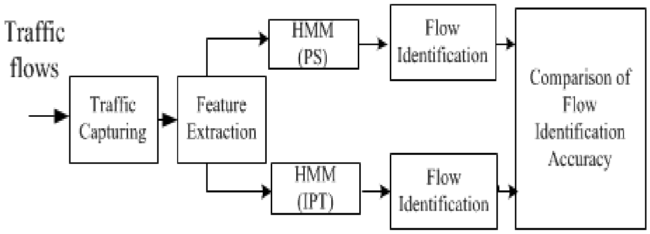

Fig. 1. Classifier block diagram.

TABLE II. Considered Traffic Statistics

App	IPT (dBμ) (mean)	IPT (dBμ) (std. dev)	PS (B) (mean)	PS (B) (std. dev)
YouTube	33	5	89	57
Email	31	11	210	343
Skype	32	10	93	159
Game	31	6	134	241

A. Modeling and Mathematical Framework

HMM is composed of hidden state variables, $x[m]$ = $[s1, s2, ., sm]$, and two dimensional observable variables, $x[m]$ = $(v1[l], v2[l])^T$, where m is representing the number of states in HMM and s is representing hidden traffic class as a state, $v1[l]$ and $v2[l]$ are inter packet time and packet size respectively. IPT can be calculated by using (2).

$$v_1[l] = 10\log(\frac{IPT}{1\mu \sec s}); \tag{2}$$

where, $v_2[l]$=packet size of mth packet

IPT and PS were assumed to be statistically independent variables. The conditional probability density functions (pdf's) for inter packet time and size are given in (3) and (4).

$$f_i^{(t)}(v_1) = \frac{(v_1 / \mathrm{w_i}^{(t)})^{(\mathrm{g_i}^{(t)}-1)} \exp(-(\mathrm{v_1} / w_\mathrm{i}^{(t)}))}{w_\mathrm{i}^{(t)}\Gamma(\mathrm{g_i}^{(t)})} (v1 > 0) \tag{3}$$

$$f_i^{(p)}(v_2) = \frac{(v_2 / \mathrm{w_i}^{(p)})^{(\mathrm{g_i}^{(p)}-1)} \exp(-(\mathrm{v_2} / w_\mathrm{i}^{(p)}))}{w_\mathrm{i}^{(p)}\Gamma(\mathrm{g_i}^{(p)})} (v2 > 0) \tag{4}$$

Where, $v_1 = v_1[1], v_1[2], \cdot \cdot \cdot , v_1[L]$ is representing IPT values, and $v_2 = v_2[1], v_2[2], \cdot \cdot \cdot , v_2[L]$ is representing PS values. The Forward variable α and the backward variable β were computed using Forward-Backward algorithm [20]. These variables are mentioned in (5) and (6).

$$\alpha_j[l] = \sum_{i=1}^{K} \mathrm{a}_j[l-1]\mathrm{A_{i},_{j}}\ f_j^t v_1[l] f_j^P v_2[l] \tag{5}$$

$$\beta_i[l] = \sum_{j=1}^{K} A_{i,j}\ f_j^t v_1[l+1]\ f_j^P v_2[l+1]\beta_j[l+1] \tag{6}$$

The likelihood for Inter packet time and packet size were computed by using (7), which is given as under

$$\lambda = P(Y / F) = \sum_{i=1}^{K} \alpha_i[l]\beta_i[l] \tag{7}$$

Test traffic was generated from known sources of YouTube, email, Skype and online game. The overall traffic in terms of bytes collected is shown in Table 3.

Delay was calculated both for PS and IPT traffic flows and their group delay for trained and training data is shown in Fig. 2. It shows that initially there was considerable delay (gap) between trained and training data but after eight iterations both started matching each other.

TABLE III. TOP FOUR APPLICATIONS TRAFFIC

Application	Amount [MB]	% of total traffic	No of flows
YouTube	548692	69	8156202
Email	52365	9	125425
Skype	25436	10	256354
Game	125425	9	354875

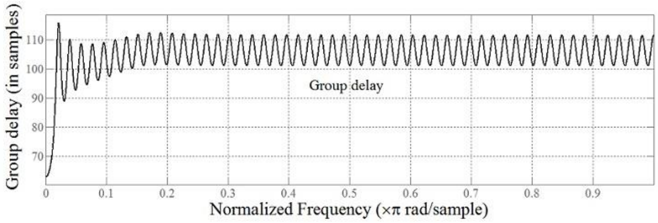

Fig. 2. Training and trained data set group delay.

PS and IPT probability density functions of these four set of traffic flows (YouTube, email, Skype & game) are shown in Fig. 3 to 6.

Fig. 3 shows that YouTube average packet size is 90 bytes and its IPT is 32 bytes. Variance between PS and IPT validates that they are two independent data sets.

As compared to YouTube traffic Fig. 4 shows that email average packet size is 200 bytes and its IPT mean is almost in the same range as that of YouTube traffic, i.e. 32 bytes. Variance between PS and IPT validates that they are two independent data sets.

Fig. 5 also validates the same variation between PS and IPT values for Skype traffic as it was observed for YouTube and email. Fig. 6 shows the variation for online game traffic.

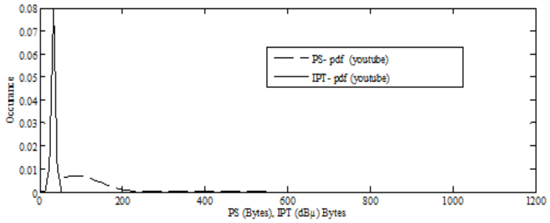

Fig. 3. PS, IPT Pdf's of YouTube.

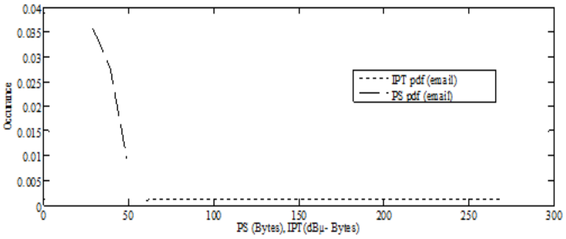

Fig. 4. PS, IPT Pdf's of email.

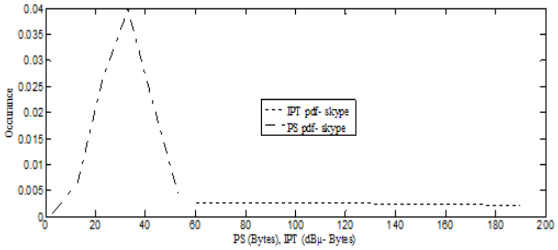

Fig. 5. PS, IPT Pdf's of Skype.

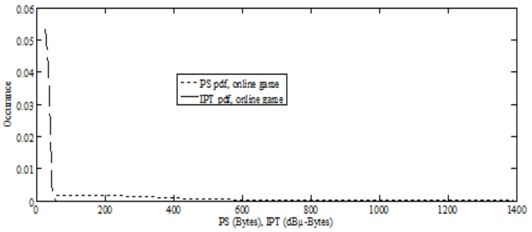

Fig. 6. PS, IPT Pdf's of online game.

TABLE IV. TRAINING SET STATISTICS

Application	IPT (dB*μ*) (mean)	IPT (dB*μ*) (std. dev.)	PS (B) (mean)	PS (B) (std. dev.)
YouTube	37	8	107	104
Email	38	9	95	68
Skype	45	9	67	25
Game	49	8	236	415

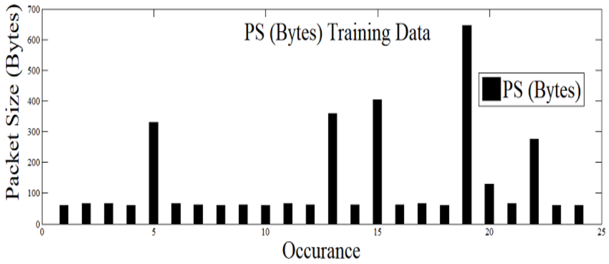

Fig. 7. PS (packet size) training data.

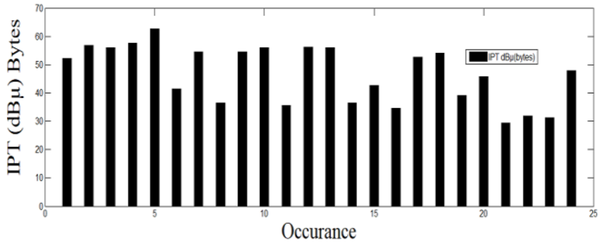

Fig. 8. IPT (inter packet time) training data.

For these models, the training data statistics are shown in Table 4. It comprises of 945 initial packets of YouTube, email, Skype and online game.

The bar plot of training data of PS and IPT is shown in Fig. 7 and 8. PS mean value of these applications is almost double as compared to the mean value of IPT. Similarly, PS and IPT standard deviation validates the distinct nature of PS and IPT traffic data.

V. ESTIMATING FLOW PARAMETERS AND RESULTS

For estimating HMM parameters, Baum-Welch introduced an iterative algorithm [21], which kept refining HMM parameters (π, A, B) until it converges to a local minimum. The Baum-Welch algorithm seeks to optimize λ via an auxiliary function $\lambda^t = (\pi^t, A^t, B^t)$, which satisfies either $\lambda = \lambda$ or $P(O/\lambda) < P(O/\lambda^t)$. It is also represented in below equation:

$$Q(\lambda`, \lambda) = \sum_q P(O, q \mid \lambda`) \log P(O, q \mid \lambda) \tag{8}$$

$Q(\lambda^t, \lambda)$ will converge to a local optimal solution, provided that the below condition is fulfilled.

$$Q(\lambda`, \lambda) \geq Q(\lambda`, \lambda`) \Rightarrow P(O \mid \lambda`) \geq P(O \mid \lambda) \tag{9}$$

$P(O/\lambda)$ give in (10) yields the results in terms of HMM parameters.

$$P(O, q \mid \lambda) = \pi_{qo} \prod_{t=1}^{T} \alpha_{qt} - {}_{qt} \prod_{K=1}^{M} b_{qtk}(O_t k) \tag{10}$$

Independent maximization of Baum auxiliary functions [21] yields a new set of model parameters given in (11), (12), and (13).

$$\overline{\pi}_i = \frac{P(O, q_0 = i \mid \lambda)}{P(O \mid \lambda)} \tag{11}$$

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T} P(O, q_{t-1} = 1, q_i = j \mid \lambda)}{\sum_{t=1}^{T} P(O, q_{t-1} = i \mid \lambda)} \tag{12}$$

$$\overline{b}_{ik}(x) = \frac{\sum_{t=1}^{T} P(O, q_t = j \mid \lambda) P(O_{ik} = x)}{\sum_{t=1}^{T} P(O, q_t = i \mid \lambda)} \tag{13}$$

By using α, β, above equations were rewritten as under in (14) whereas parameters re-estimation in terms of α and β is as under in (15).

$$P(O, q_t = i \mid \lambda) = \alpha_t(i)\beta_t(i), P(O/\lambda) = \sum_{i=1}^{N} a_t(i)\beta_t(i) \tag{14}$$

$$\overline{\pi}_i = \frac{\alpha_0(i)\beta_0(i)}{\sum_{j=1}^{N} \alpha_0(j)\beta_0(j)} \tag{15}$$

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T} \alpha_{t-1}(i) a_{ij} \beta_t(j) \sum_{k=1}^{M} b_{jk}(O_{tk})}{\sum_{j=1}^{N} \alpha_{t-1}(i) \beta_{t-1}(i)} \tag{16}$$

$$\overline{b}_{ik}(x) = \frac{\sum_{t-1}^{T} \alpha_t(i) \beta_t(i) b_{ik}(O_t k)}{\sum_{t=1}^{T} \alpha_t(i) \beta_t(i)} \tag{17}$$

The above equations represent new sets of estimated parameters learned with the help of Expectation Maximization algorithm.

A. Traffic Flows Estimation

HMM Viterbi was applied to find out the most likely path for the hidden Markov model as specified by the state transition matrix (A), and emission matrix (B). Model parameters were iteratively improved by using Viterbi Algorithm. PS and IPT states state transition as shown in Fig. 9 and 10 were used to optimize likelihood of each state.

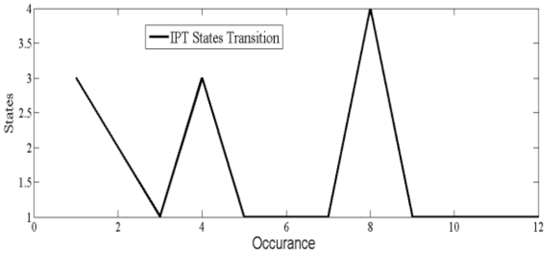

Fig. 9. States transition (IPT).

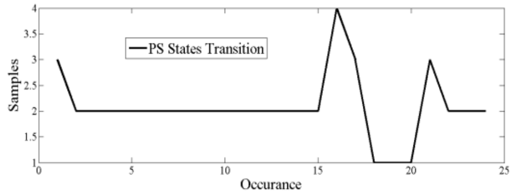

Fig. 10. States transition (PS).

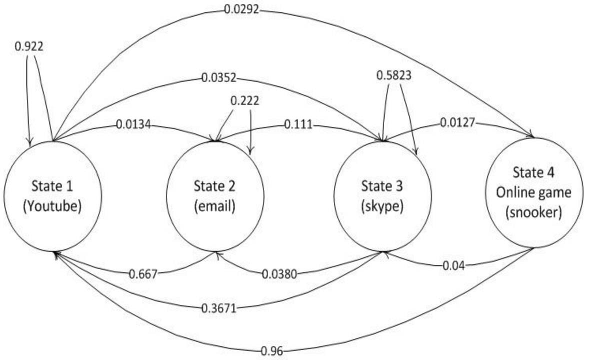

Fig. 11. State transition diagram (PS).

These figures indicate that YouTube was the mostly found state both for PS and IPT transitions. It also matches with actual traffic which was generated from different traffic sources. After computing YouTube, email, Skype and online game, their state transition probabilities are shown in Fig. 11. It shows that YouTube has 92.2% probability to stay within the same state, which reflects that overall traffic is mostly dominated by YouTube, and there are higher chances that the YouTube state probability always remains very high as compared to other flows.

The traffic flow identification accuracy results for PS and IPT are shown in Table 5. The traffic identification accuracy PS was up to 92%, whereas for IPT, the achieved accuracy was 87%.

The modeling results of YouTube, email, Skype and game for PS and IPT are shown in Table 6 and 7. The results are shown through a confusion matrix. All correct classification were shown italic in below tables.

TABLE V. PS AND IPT ACHIEVED ACCURACY COMPARISON

No of Packets (training data)	Accuracy Achieved (PS)	Accuracy Achieved (IPT)
5	80%	60%
12	91.67%	75%
946	92%	87%

TABLE VI. CLASSIFICATION RESULTS CONFUSION MATRIX (PS)

Application	YouTube	Email	Skype	Game
YouTube	*91.93%*	6.5%	1.2%	0.37%
Email	5.4%	*84.20%*	1.54%	8.86%
Skype	3.54%	5.48%	*81.25%*	9.73%
Game	4.58%	6.54%	9.34%	*79.54%*

TABLE VII. CLASSIFICATION RESULTS CONFUSION MATRIX (IPT)

Application	YouTube	Email	Skype	Game
YouTube	*81.00%*	9.0%	3.0%	7.0%
Email	8.0%	*73.20%*	8.00%	10.80%
Skype	9.00%	8.00%	*69.00%*	14.00%
Game	8.00%	9.00%	11.00%	*72.00%*

Row 1 in Table 6 shows that for YouTube application achieved accuracy for PS based modeling was 91.93%, whereas 6.5% of the YouTube traffic had been classified as email, 1.2% as Skype, and 0.37% had been classified as online game. This shows that accuracy of classifier was up to 91.93% for YouTube, 84.20% for email, 81.25% for Skype and 79.54% for online game. Similarly, in case of IPT, the Table 5 shows that accuracy of classifier for YouTube traffic was 81%, for email 73.2%, for Skype it was 69% and for online game, it was 72%. For different flows, we considered traffic in one direction only and that could be one of the reasons that to a certain extent, the accuracy was 69% for Skype. Considering traffic in both directions may improve the accuracy.

VI. CONCLUSION

With rapid advancements in Internet of Things, the network resources are no more unlimited, and bandwidth hungry multimedia applications are consuming the major part of available bandwidth. Traffic classification is key to network security solution and management architectures [22], [23]. In this paper, a novel HMM based modeling technique has been proposed that can classify internet traffic based on their statistical properties, i.e. PS and IPT. The traffic classifications have been done by using minimum number of statistical parameters, which reduced computational complexity and overall load on network systems. The comparative analysis of PS and IPT shows that achieved classification accuracy for PS based model was 92% and for IPT it was 81%. The achieved accuracy suggests that proposed modeling framework can be part of a multi traffic classifier system. Moreover, PS and IPT combination could also result in better accuracy and can be an area of future work on traffic classification.

Acknowledgment

Authors are grateful to Pakistan Telecommunication Company Limited and Post graduate Laboratory of Electrical Engineering Department, University of Engineering and Technology, Lahore for allowing the use of their resources for the verification of results and concluding the research.

References

[1] Karagiannis, Thomas, Papagiannaki, Konstantina, Faloutsos and Michalis.BLINC: multilevel traffic classification in the dark.ACM SIGCOMM Computer Communication Review,ACM,2005,35(4):229-240.

[2] Kim, Hyunchul, Claffy, Kimberly C, Fomenkov, Marina, Barman, Dhiman, Faloutsos, Michalis and Lee, KiYoung. Internet traffic classification demystified: myths, caveats, and the best practices. Proceedings of the 2008 ACM CoNEXT conference,ACM,2008.11

[3] Wu, Yulei, Min, Geyong, Li, Keqiu, Javadi and Bahman.Performance analysis of communication networks in multi-cluster systems under bursty traffic with communication locality.Global Telecommunications Conference, GLOBECOM, IEEE, 2009,1-6

[4] Lim, Yeon-sup, Kim, Hyun-chul, Jeong, Jiwoong and Kim, Chongkwon and Kwon, Ted Taekyoung and Choi, Yanghee. Internet traffic classification demystified: on the sources of the discriminative power. Proceedings of the 6th International Conference, ACM,2010,9

[5] Zhang, Jun and Chen, Chao and Xiang, Yang and Zhou, Wanlei and Xiang, Yong. Internet traffic classification by aggregating correlated naive bayes predictions. IEEE Transactions on Information Forensics and Security. IEEE. 2005. 8(1):5-15

[6] Roughan, Matthew and Sen, Subhabrata and Spatscheck, Oliver and Duffield, Nick. Class-of-service mapping for QoS: a statistical signature-based approach to IP traffic classification.Proceedings of the 4th ACM SIGCOMM conference on Internet measurement, ACM,2004, 135-148

[7] Moore, Andrew W and Papagiannaki, Konstantina. Toward the accurate identification of network applications, International Workshop on Passive and Active Network Measurement,Springer, 2005.41-54

[8] Kim, Hyunchul and Claffy, Kimberly C and Fomenkov, Marina and Barman, Dhiman and Faloutsos, Michalis and Lee, KiYoung. Internet traffic classification demystified: myths, caveats, and the best practices,ACM,2008:11

[9] Leland, Will E and Willinger, Walter and Taqqu, Murad S and Wilson, Daniel V. On the self-similar nature of ethernet traffic, ACM SIGCOMM Computer Communication Review, ACM, 1995. 25(1):202-213

[10] Domanski, Adam,Domanska, Joanna and Czachorski, Tadeusz. International Conference on Next Generation Wired and Wireless Networking,Springer,2008:156-168

[11] Nguyen, Thuy TT and Armitage, Grenville. A survey of techniques for internet traffic classification using machine learning,IEEE Communications Surveys & Tutorials. IEEE,2008. 10(4):56-76

[12] Moore, Andrew W and Zuev, Denis. Internet traffic classification using bayesian analysis techniques,ACM SIGMETRICS Performance Evaluation Review,ACM, 2005. 33(1):50-60

[13] Zander, Sebastian and Nguyen, Thuy and Armitage, Grenville. Automated traffic classification and application identification using machine learning,The IEEE Conference on Local Computer Networks 30th Aniversary (LCN'05) l, IEEE. 2005.250-257

[14] Nguyen, Thuy TT and Armitage, Grenville and Branch, Philip and Zander, Sebastian. Timely and continuous machine-learning-based classification for interactive IP traffic, IEEE/ACM Transactions on Networking (TON), IEEE/ACM. 2012. 20(6):1880-1894

[15] Dainotti, Alberto and De Donato, Walter and Pescape, Antonio and Rossi, Pierluigi Salvo.Classification of network traffic via packet-level hidden Markov models. IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference, IEEE. 2008.1-5

[16] Gao, Haihong and Best, Tish and Pendse, Ravi and Sawan, Mahmoud Edwin. Traffic classification and observer design of cable networks, IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR). 2012.1-6

[17] Szabó, Géza and Orincsay, Dániel and Malomsoky, Szabolcs and Szabó, István. On the validation of traffic classification algorithms, International Conference on Passive and Active Network Measurement,Springer.2008.72-81

[18] Park, Byung-Chul and Won, Young J and Kim, Myung-Sup and Hong, James W. Towards automated application signature generation for traffic identification,NOMS 2008-2008 IEEE Network Operations and Management Symposium,IEEE. 2008. 160-167

[19] Dainotti, Alberto and Pescape, Antonio and Rossi, Pierluigi Salvo Salvo and Iannello, Giulio and Palmieri, Francesco and Ventre, Giorgio.Qrp07-2: An HMM approach to internet traffic modeling,IEEE Globecom 2006, IEEE. 2006.1-6

[20] Rabiner, Lawrence R. A tutorial on hidden Markov models and selected applications in speech recognition,Proceedings of the IEEE. 1989. 77(2):257-286

[21] Baum, Leonard E and Petrie, Ted and Soules, George and Weiss, Norman.A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains,The annals of mathematical statistics, JSTOR. 1970. 41(1):164-171

[22] Jiang, Hongbo and Moore, Andrew W and Ge, Zihui and Jin, Shudong and Wang, Jia. Lightweight application classification for network management,Proceedings of the 2007 SIGCOMM workshop on Internet network management,ACM. 2007.299-304

[23] Marques, Oge and Baillargeon, Pierre. Design of a multimedia traffic classifier for Snort,Information management & computer security,Emerald Group Publishing Limited. 2007. 15(3):241-256

[24] Roesch, Snort - lightweight intrusion detection for networks. In Proceedings of USENIX LISA99, 1999

[25] Qazi, C Tu, Chiang, Miao, Sekar, and Yu, Simple-fying middlebox policy enforcement using sdn,ACM SIGCOMM computer communication review, 2013.43(4):2738

[26] Dainotti, Pescap, Salvo Rossi, Palmieri, Ventre, Internet Traffic Modeling by means of Hidden Markov Models; Computer Networks (Elsevier), 2008.52(1)4:2645-2662

[27] Postel, User datagram protocol. internet engineering task force, tech. rep.,1980. RFC 768.

[28] Touch, Kojo, Lear, Mankin, Ono, Stiemerling, and Eggert, Service name and transport protocol port number registry, The Internet Assigned Numbers Authority (IANA), 2013.

[29] Cohen, Incentives build robustness in bittorrent, in Workshop on Economics of Peer-to-Peer systems,2003.6:6872

[30] Moore and Papagiannaki, Toward the accurate identication of network applications, in International Workshop on Passive and Active Network Measurement,springer, 2005:4154

[31] Roughan, Sen, Spatscheck, and Dueld, Class-of-service mapping for qos: a statistical signature-based approach to ip trac classication, in Proceedings of the 4th ACM SIGCOMM conference on Internet measurement,ACM, 2004:135148

[32] Madhukar and Williamson, A longitudinal study of p2p trac classication, in 14th IEEE International Symposium on Modeling, Analysis, and Simulation, IEEE,2006:179188