Jurnal Teknologi

Full paper

Internet Traffic Classification Algorithm Based on Hybrid Classifiers to Identify Online Games Traffic

Hamza Awad Hamza Ibrahim[a*], Sulaiman Mohd Nor[a], Ali Ahmed[b]

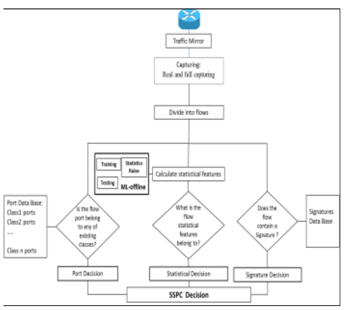
[a]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author: hamysra76@hotmail.com

Article history

Received :12 July 2012
Received in revised form:
4 April 2013
Accepted :15 April 2013

Graphical abstract



Abstract

Classification of interactive applications such as online games has gained more attention in the last few years. However, most of the current classification methods were only valid for offline classification. The three common classification methods i.e. port, payload and statistics based have some limitations. This paper exploits the advantages of all the three methods by combining them to produce a new classification algorithm called SSPC (Signature Static Port Classifier). In the proposed algorithm, each of the three classifiers will individually classify the same traffic flow. Based on some priority rules, SSPC makes classification decision for each flow. The SSPC algorithm was used to classify online game (LOL) traffic in two stages, initially offline and later online. SSPC produces a higher accuracy of 91% on average for online classification when compared with other classifiers. In addition, as demonstrated in the real time online experiments done, SSPC algorithm uses a short time to classify traffic and thus it is suitable to be used for online classification.

Keywords: Internet traffic classification; online games; online classification; machine learning; classification algorithm

Abstrak

Sejak beberapa tahun lepas, perhatian yang lebih telah diberikan pada pengelasan aplikasi interaktif seperti permainan atas talian. Kebanyakan daripada pengelasan semasa hanya sah untuk pengelasan luar talian(offline). Terdapat kekurangan terhadap tiga kaedah klasifikasi lazim yang berasas pengkalan, muatan dan statistik. Kertas kerja ini mengeksploitasi kelebihan tiga kaedah tersebut dan menggabungkan ketiga bagi menghasilkan satu algoritma pengelasan baru dikenali sebagai SSPC (Signature Static Port Classifier). Melalui algoritma yang dicadangkan, setiap dari tiga pengelas akan mengelaskan aliran trafik sama secara berasingan. SSPC membuat keputusan pengelasan untuk setiap aliran berdasarkan peraturan yang menentukan keutamaan. Algoritma SSPC diguna untuk mengelaskan trafik permainan atas talian (LOL) dalam dua peringkat, asalnya luar talian dan kemudian atas talian. SSPC menghasilkan ketepatan lebih tinggi iaitu 91% secara purata apabila dibandingkan dengan pengelas-pengelas lain. Sebagai tambahan, sepertimana yang ditunjukkan dalam ujikaji masa nyata atas talian yang dijalankan, algoritma SSPC menghasilkan tempoh masa singkat untuk mengklasifikasi trafik. Ini menunjukkan SSPC amat bersesuaian diguna untuk pengelasan atas talian.

Kata kunci: Pengelasan trafik internet; permainan atas talian; pengelasan atas talian; pembelajaran mesin; algoritma pengelasan

1.0 INTRODUCTION

Network managers are usually interested to know the traffic carried in their networks for the purposes of optimizing network performance and security issues. Therefore, Internet traffic classification is something important, particularly interactive traffic applications such as online games and Skype.

Port classification assumes that applications used well-known port and the classifier used this port number to determine the application type. However, most Internet applications used unknown port number or more than one application used the same port number, which indicates the failure of port base classification [1]. Another classification method is payload based (deep packet inspection), which is individual packet inspection looking for specific signatures. However, using of this technique faced by two problems; first, it difficult to detect non-standard port by using packet inspection, because these packets are encrypted. Second, deep packet inspection touches users' privacy. Machine Learning

(ML) technique was appeared to solve the problem of past classification methods. ML [2] [3] used artificial intelligence to classify IP traffic, which provide a suitable solution by extracting right information from application features [4]. Moreover, some of the ML algorithms are suitable for Internet traffic flow classification at a high speed [5]. Because of the rapid sort of real time applications, the main issue is the time of collecting the statistical values (rules building), which assumed to be extremely short. Most of the proposed ML classification methods was limited for offline traffic classification and cannot support online classification [6]. Online classification means the decision of what is flow/packet belong to; assume to be on the time of capturing. Such like any hardware classifier (PacketShaper, SANGFOR) installed on network router, which is classifying with the passage of the traffic.

The researchers look for online game traffic classification from two different points of view. The first group [7] [8] [9] looks to help game users by enhancing Quality of Service (QoS) present in the network. This means identify game traffic and then give it a high bandwidth priority. The second group [10], which are adverse of the first group, which classified online game to managing Internet traffic and decreases bandwidth consumed by online game users. Other researchers look to game traffic as a part of network traffic, then they aim to identify total Internet traffic. In any case, online game traffic classification is an essential issue with the increasing in games traffic.

Most previous literatures [6] [11] [12] [13] [14] [15] [16] [17] provide a classifiers work with real time traffic, but few of them [18] provided classifiers can get an online decision. Identification of online game is not easy because of several reasons; first, most of the online games are P2P applications. Second, games traffic do not share similar characteristics [19].

This paper describes the development of an online Signature Statistic Port Classifier (SSPC) algorithm, which can identify Internet traffic shortly after traffic is captured. The classifier differs from other works since it takes the classification decision based on three parallel different methods. League of Legends (LOL) is the most widely online game used in the campus network, which belong to Garena Plus online games. SSPC algorithm was used to classify LOL against the other applications.

Section 2 describes and analyses the related works. General concepts of classification mechanisms, the three partial algorithms, and SSPC algorithm are discussed in Section 3. In section 4, the experiments and analyzed results are shown. Finally, the conclusion is provided in section 5.

■2.0 RELATED WORK

This section addresses the related works from two points of view; the articles of online classification methods and articles of online games classification methods.

In [7] the authors proposed Automated Network Games Enhancement Layer (ANGEL), which is a system used to improve QoS of online games. Most of the system parts are located on the ISP side. ANGEL captures a copy of the traffic between ISP and customer networks then classifies it into applications type (games and others). Then, it forwards the priority to the customer. However, the authors' does not discuss how the classification was done and what are the features used with ML. In addition, ISP has thousands of customer's networks and it is difficult to cover all of them.

The authors of [11] focus on traffic measurement in high speed network. The paper analyzes Internet applications to see the traffic measurement characteristics. Then design flow measurement system of high speed network based on Linux kernel is described. The system was built over some methods; firstly, from a perspective of Network Interface Card (NIC); the system designed a Hash function (group of rules) to classify packet processed by 32-bit systems instead of interrupt to communicate with OS. Second, the system identify the new P2P service by calculating key hash value if there are no existing matching rules. Some shortcoming was observed such as the authors do not provide details what features are used to build hash rules. As well, the system defines any new traffic flow as new P2P. Moreover, the paper relies heavily on port numbers to identify traditional applications.

The paper [12] proposed a dynamic online method to classify Internet traffic. The method used two levels: overall traffic level and application level. Data mining algorithms are used to continue updated considered datasets. The proposed method has three parts: i) Traffic model; which is: preparing the dataset, selecting the features, and updating the model for the case of new application. ii) Traffic classification; to classify traffic based on the gained features. iii) Change detection; which is run periodically to check if there is a new application. While the paper title includes the words "online traffic classification", but there is no online classification. Also no details about traffic features used for classification are provided.

The study [13] proposed an approach for online classification for TCP traffic based on the first n packets. The approach used information from the first n packets to decide which kind of application the flow belongs to. The authors used correlation-based feature selection (CFS) [20] to select optimum features. However, it is something untrusted to classify flow include thousands packets based on the first few packets. This is because the first packets in many flows can differ statistically from the rest of the packets. Moreover, the paper did not provide details how the online decision was taken.

In [18], the authors proposed a network processor (NP) classifier, which is based on online hybrid traffic to identify P2P traffic. The classifier is based on two stages: hardware static characteristics and software Flexible Neural Tree (FNT) [21]. In the first stage, the hardware classifier (based on payload and port) filters P2P traffic. In the second stage, the software classifier (based on ML statistical features) is used as statistical diction maker. However, the classifier depends on hardware (NP), which is an additional cost.

In [14], traffic classifier based on Support Vector Machine (SVM) was presented. The dataset include three traces collected from three different places. Based on statistical features, the classifier used the first ten packets to identify the flow. As in the previous paper, while the paper title content the words "online classification", but there is no online decision. In addition, how to classify flow includes a huge amount of packets based only on ten packets.

The researchers of [15] proposed a wireless mesh network traffic classification using C4.5. The dataset was represented by sub-flow and application behaviors. Based on the statistical features of the first n packets, the classifier clusters the flow to one of the defined applications. Similar to the previous work, the datasets were captured in real time; however, there is no online classification.

[17] is a flattering work which proposes a method suitable for identifying the application associated with a TCP flow. This method based on total data length sent by client (ACK-Len ab) or server (ACK-Len ba) before it received ACK packets. The work analysis TCP flows to get the characteristics of the two adopted features (ACK-Len ab and ACK-Len ba). The proposed method was verified by using ML classifier (C4.5) to classify four types of Internet application (WWW, FTP, EMAIL and P2P). In the same manner, no online classification was provided.

[16] develop a classifier which quickly identifies an application at any point of a flow's lifetime. Thus, the ML classifier was trained by using sets of features calculated from multiple sub-flows at different points. The classifier recognizes the flow either way (forward or backward) by features swapped called Synthetic Sub-flow Pairs (SSP). Assistance of Clustering Techniques (ACT) as unsupervised clustering ML technique was used to automate the selection process. The problem is different datasets from different dates (may be different network) was considered for ML. This is not consistent with the rule of similarity of training and testing datasets in a network environment.

[22] is recent work proposes multistage classifier. Binary Particle Swarm Optimization (BPSO) is a method applied by this work to select the best flow features. Three methods (port, payload, and statistical based) were integrated into the multistage classifier. The idea of this work is extremely good; however it was not tested for online classification which identifies the traffic with capture speed. In addition, classifier can only make decision based on the first stage (port based method).

■3.0 ONLINE INTERNET TRAFFIC CLASSIFICATION MECHANISM

3.1 General Concepts and Definitions

Definition 1: Flow is a group of packets share the same 5-tuples (source address, destination address, source port, destination port, and transport protocol). Flow can be represented by TCP or UDP packets. We consider unidirectional flows, which defines client server traffic as different from server client traffic. Definition 2: Real time traffic is the Internet traffic captured from the campus network during the period of experiments. Definition 3: Offline decision is the decision by the classifier about the flows identification, which is taken offline after capturing time. Definition 4: Online decision is the decision by the classifier about the flows identification, which is taken online within capturing time. Since the Internet applications are continuously being developed, it is difficult to classify the traffic by using only one classification method [22]. This paper develops online Signature Statistical Port Classifier (SSPC), which is making classification decision near to the capturing time. The classifier makes his final decision based on three parallel partial decisions (port classifier, signature classifier, and statistic classifier)

Port Classifier

As mentioned in section 1, port based classification cannot achieve a high accuracy all the time. In this paper, port classification is an algorithm using as a part of the classification system and it represent low priority of SSPC classification decision. In most cases, SSPC classification decisions not making based alone on port classifier, but it shares the decision with the other two classifiers. We develop port classifier algorithm as a part of SSPC algorithm. Port classifier makes his own decision based on port DataBase. Easley, the port classifier algorithm compares the port number of the flow with the ports DataBase. If found then the flow will classify based on port classifier rules.

Signature Classifier

Payload classification can achieve high accuracy, but it cannot work with encrypted traffic. As before, SSPC did not fully depend on payload, but it does only represent a part of the final decision. We develop signature classifier algorithm which is the second part of SSPC algorithm; this algorithm take classification decision based on some saved signatures. We add some general signatures (such as DNS query and http host) for the considered applications, which are extracting from the application layer. If the flow has a signature from the signature data base, it will classify based on what signature belong to.

Statistic Classifier

The main problem meats ML classification is the high false positive. To reduce this problem we consider two issues; firstly, the offline training datasets were continuously updated and collected manually from the same network we need to classify. Secondly, statistic classifier was supported by the other two classifiers. Statistic classifier algorithm is the third part of SSPC algorithm. Also as before, ML classifier represents a part of the system decision.

SSPC

In the purpose of increasing the classification efficiency, SSPC is proposed. SSPC is a result of the three previous classifiers decisions. Differing from the previous works [18] and [22], SSPC did not base on hardware part; also he did not take his decision based only on one method. The online flow classification was occurred after comparison of three stage classification decisions. Moreover, SSPC was tested for online classification decision.

3.2 SSPC Architecture

Figure 1 illustrate the classifier stages, which started by fully packets capturing using traffic mirror. Before delivered to the three classifiers, the traffic was divided into flows based on the 5-tuples. Each flow will classify three times by each of the three classifiers. The port classifier compares the captured flow ports with a list of saved port numbers. If the captured flow belongs to any group of saved ports, it will identify as its group as. The second classifier (statistical) works parallel with the first classifier. Based on offline training and testing datasets, some classification rules were building. Based on these rules, the statistical classifier (algorithm) makes his online decisions to identify the captured flows. On the other hand, the signature classifier will classify the same traffic at the same time of the previous two classifiers; the classifier will compare a part of captured flow with signature data base. If the signature matches any of saved signatures, the classifier will take his online decisions to identify the captured flows. SSPC is an algorithm which compares between the three classifiers result and makes his online classification decision based on some priorities rules.

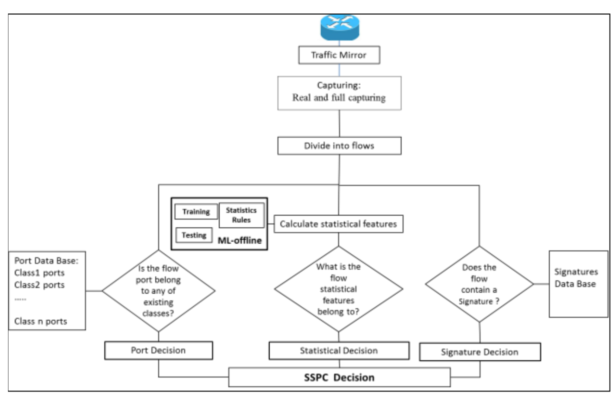


Figure 1 SSPC architecture

3.3 SSPC Algorithm

The SSPC algorithm is shown in algorithm 1. Matlab version 7.5.0.342 (R2007b) was used to develop the algorithm. The SSPC algorithm consists of three partial classifier algorithms (described in section 4.2); each classifier has it's own classification decision. Because of the accuracy of the signature classifier, the first priority in SSPC decision making goes to the signature classifier. If the signature classifier makes any decision about this flow, then SSPC final decision will to the decision made by the signature classifier. The second priority of SSPC happen in the case when of all the partial classifiers cannot make any decision about this flow. In this case, SSPC will classify the flow as unknown. The third priority of SSPC occurs in the case when the statistics and the port classifiers have the same decision and signature classifier have no decision (unknown). In this case, SSPC identifies this flow based on the statistic and port classifiers. When the statistic and port classifiers have different opinions about the flow, SSPC will classify this flow as the statistic classifier as. The SSPC decision is based on the port classifier in only one case. This occurs when the port classifier has a decision whilst both statistic and signature classifiers have no decision about the flow.

The pseudocode for SSPC algorithm is as follows:

```
// Define variables
Array port_DataBase;
Array signatures_DataBase;
string statistical_rules;
start packets capturing;
//divide captured packets into flows
if the packet belongs to an existing flow
then adds this packet to the existing flow
else
initialize a new flow;
end if
//run the three algorithm classifiers in parallel
calculate flow statistic_values;
update statistic_rules;
for (flow=0; flow=-1;flow++)
{
// check signature
inspects n packets in the flow;
if class_signatures found
decision1=classify this flow according to
signatures_DataBase;
else
decision1=classify this flow as unknown;
end if
//check statistical
if statistical_of_the_flow achieved any statistic_rules
decision2=classify this flow according to statistic_rules;
else
decision2=classify this flow as unknown;
end if
// Check port
if flow_port in port_DataBase
decision3=classify this flow according to port_DataBase;
else
decision3=classify this flow as unknown;
end if
end if
// calculate SSPC decision
if decision1 != "unknown"
SSPC_decision= decision1;
else if decision1 = decision2= decision3="unknown"
SSPC_decision= ="unknown"
else if decision2 = decision3
SSPC_decision= decision2;
else if decision1 = decision2= "unknown"
SSPC_decision= decision3;
else
SSPC_decision = decision2;
end ;
}
```

Algorithm1 SSPC algorithm

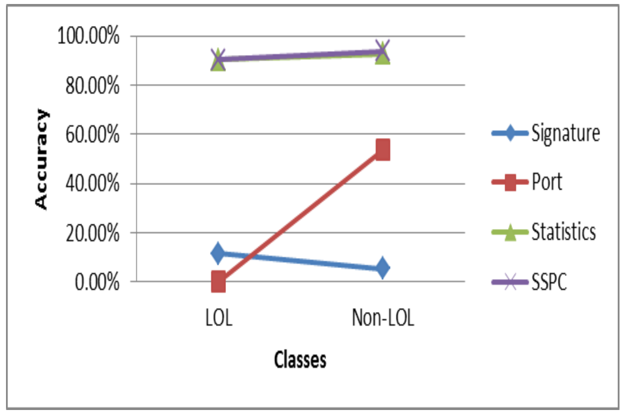


Figure 2 Classifiers offline accuracy

4.0 EXPERIMENTS AND ANALYSIS

In order to evaluate the methodology, different datasets composed of thousands of flows were considered. League of Legends (LOL) is the most widely online game used in UTM campus network, which belong to Garena Plus online games. LOL represent more than 22% of campus online games. Real time LOL and non-LOL traffic were collected from the campus network. Table 1 show the flows considered for both classes, which are obtained manually through the monitored clients (IPs). By this way, we ensure the training and testing datasets were collected from the same network without the need for standard (labeled) datasets. LOL traffic was generated by playing the actual game online using both s wireless and cables network. Non-LOL traffic includes http, https, FTP control, FTP data, and Skype, which is real time traffic collected in the same way as collecting the LOL data.

Table 1 Considered flows for offline classification

Class (applications)	considered flows
LOL	52
Non-LOL	4205

For ML training purpose, we captured traffic from some monitored clients. Offline ML classification was performed to select the optimum features and algorithm. After some filtering, Rule.PART algorithm within Weka [23] was selected as the ML classifier. Rule.PART rules were inserted into the statistic classifier algorithm. Interarrival time and packets length (size) were used as traffic features. From these two features, ten statistic factors were calculated are shown in Table 2.

Table 2 Selected features

Max of Interarrival time
Min of Interarrival time
Mean of Interarrival time
Variance of Interarrival time
Standard deviation of Interarrival time
Max of packet length
Min of Packet length
Mean of Packet length
Variance of Packet length
Standard deviation of Packet length

Before going into online decisions experiments, offline works were performed to validate the methodology. SSPC was used to distinguish between LOL and non LOL traffic. First; each classifier (port, signature and statistical) was used over each class dataset (table 1) separately. Second; in the same manner, SSPC algorithm was used over each class dataset. Table 3 and figure 2 show each individual classifier accuracy and SSPC accuracy. For both classes of the considered datasets, SSPC shows the higher accuracy compared with the other partial classifiers.

Table 3 Results of offline classification

	Signature	Port	Statistics	SSPC
LOL	11.54%	0.00%	90.38%	90.38%
Non-LOL	5.32%	53.63%	92.79%	93.73%

In the online decision, the same offline applications through two different experiments were considered. Similar as in the offline experiments; the applications were run in the monitored clients. The testing dataset generated were totally different from the training dataset. As an example for some clients, we only run LOL games and at the same time check what is the decision of each individual classifier and SSPC decision. Table 4 shows the number of flows generated by each of online experiment. Table 5 and figure 3 illustrate the accuracy of online decisions. In experiment one; SSPC results in a higher accuracy when compared with the other three classifiers. In experiment two, SSPC produces the highest accuracy compared to the other classifiers when we deal with Non-LOL classification and the same accuracy to statistical classifier when classifying LOL. The last column in table 5 shows the average classification time (in seconds) for each flow. As an example, classifying single LOL flow in experiment 1 takes 0.02 second immediately after the end of captured flow.

Table 4 number of flows for online classification

	Experiment 1	Experiment 2
LOL	51	458
Non-LOL	1393	1433

Table 5 Results of online classification

	Port	Statistics	Signature	SSPC	flow/S
Experiment 1					
LOL	0.00%	80.39%	43.14%	88.24%	0.02
Non-LOL	47.81%	81.98%	4.01%	92.76%	0.08
Experiment 2					
LOL	0.00%	94.32%	24.67%	94.32%	0.05
Non-LOL	45.24%	81.55%	4.37%	89.99%	0.07

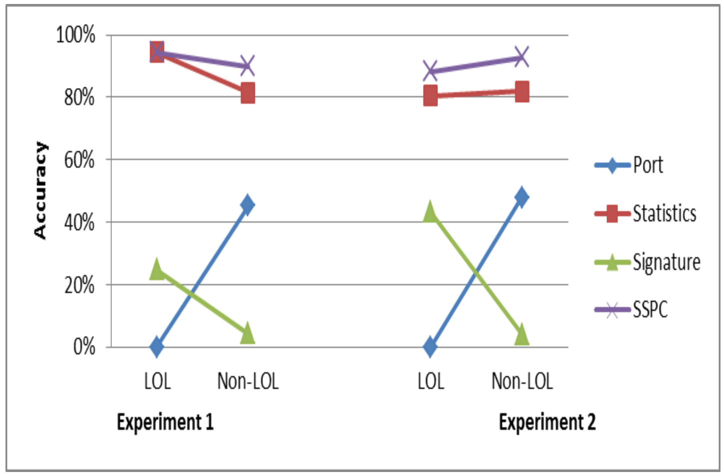


Figure 3 Classifiers online accuracy

5.0 CONCLUSION

Port based classifier has the advantage of non-complexity; however, it cannot achieve high accuracy with applications of unknown port numbers. On the other hand, payload classifiers have the advantage of accurate classification, but incapable to classify encrypted traffic. Statistic classifier has the benefit of classifying encrypted traffic, but it has the problem of high false positive. In this paper, Signature Statistical Port Classifier (SSPC) algorithm for online traffic classification is proposed. Each of the three partial classifiers (based on the three methods) makes decision about each traffic flow. SSPC algorithm, based on some priority rules calculates the final decision from the outcome of the three classifiers. The proposed algorithm was used to distinguish between LOL and non-LOL traffic. Real time datasets (more than 7600 flows) were captured from a campus environment, which includes LOL and non-LOL (http, https, FTP-data, FTP-control, and Skype). The SSPC was tested in two stages, offline and online. The results of the offline experiments show that SSPC has higher accuracy compared to the three partial classifiers. To further validate the robustness of the algorithm, online classification using SSPC algorithm was carried out. The classification decision was made immediately after the end of flow capture. The results show a higher accuracy when compared with the individual classifiers. Thus, SSPC has achieved the two objectives; the first, capitalizing on the effectiveness of the individual classifiers which sums up to effectively increase SSPC accuracy. Second, SSPC can still classify the online Internet traffic without any compromise in delay.

References

[1] Nguyen, T.T.T., Armitage, G. 2008. A Survey of Techniques for Internet Traffic Classification using Machine Learning. *Ieee Commun Surv Tut.* 10(4): 56–76. doi:Doi 10.1109/Surv.2008.080406.

[2] Jesudasan, R. N., Branch, P., But, J. 2010. Generic Attributes for Skype Identification Using Machine Learning. Technical Report 100820A.

[3] Alshammari, R., Zincir-Heywood, A. N. 2010. An Investigation on the Identification of VoIP Traffic: Case study on Gtalk and Skype. In: Network and Service Management (CNSM), 2010 International Conference on, 25-29 Oct. 2010. 310–313.

[4] Yu, J., Lee, H., Im, Y., Kim, M.S., Park, D. 2010. Real-time Classification of Internet Application Traffic using a Hierarchical Multi-class SVM. *Ksii T Internet Inf.* 4(5): 859–876. doi:DOI 10.3837/tiis.2010.10.009.

[5] Soysal, M., Schmidt, E. G. 2010. Machine Learning Algorithms for Accurate Flow-based Network Traffic Classification: Evaluation and Comparison. *Perform Evaluation.* 67(6): 451–467. doi:DOI 10.1016/j.peva.2010.01.001.

[6] Gu, R., Wang, H., Ji, Y. 2010. Early Traffic Identification Using Bayesian Networks. 564–568.

[7] But, J., Nguyen, T., Stewart, L., Williams, N., Armitage, G. 2007. Performance Analysis of the ANGEL System for Automated Control of Game Traffic Prioritisation. 123–128.

[8] But, J., Williams, N., Zander, S., Stewart, L., Armitage, G. 2006. Automated Network Games Enhancement Layer-A Proposed Architecture.

[9] But, J., Armitage, G., Stewart, L. 2008. Outsourcing Automated QoS Control of Home Routers for a Better Online Game Experience. *Ieee Commun Mag*. 46(12): 64–70.

[10] Chengjie, G., Shunyi, Z. 2010. A Novel P2P Traffic Classification Approach Using Back Propagation Neural Network. In: Communication Technology (ICCT), 2010 12th IEEE International Conference on, 11-14 Nov. 2010. 52–55.

[11] Xu, C., Tang, H., Zhao, G. F. 2008. Traflow: Design and Complementation of a Real Time Traffic Measurement System in High-Speed Networks. 2008 Ifip International Conference on Network and Parallel Computing, Proceedings. 341–344. doi:Doi 10.1109/Npc.2008.10.

[12] Xu, T., Qiong, S., Xiaohong, H., Yan, M. 2009. A Dynamic Online Traffic Classification Methodology Based on Data Stream Mining. In: Computer Science and Information Engineering, 2009 WRI World Congress on, March 31 2009-April 2, 2009. 298–302.

[13] Hong, M.-h., Gu, R.-t., Wang, H.-x., Sun, Y.-m., Ji, Y.-f. 2009. Identifying online traffic based on property of TCP flow. The Journal of China Universities of Posts and Telecommunications 16(3): 84–88. doi:http://dx.doi.org/10.1016/S1005-8885(08)60231-9.

[14] Gu, C., Zhang, S., Huang, H. 2011. Online Internet Traffic Classification Based on Proximal SVM. *Journal of Computational Information Systems*. 7(6): 2078–2086.

[15] Gu, C., Zhang, S., Xue, X., Huang, H. 2011. Online Wireless Mesh Network Traffic Classification Using Machine Learning. *Journal of Computational Information Systems*. 7(5): 1524–1532.

[16] Nguyen, T.T.T., Armitage, G., Branch, P., Zander, S. 2012. Timely and Continuous Machine-Learning-Based Classification for Interactive IP Traffic. Networking, IEEE/ACM Transactions on PP(99). 1–1. doi:10.1109/tnet.2012.2187305.

[17] Sun, M. F., Chen, J. T. 2011. Research of the Traffic Characteristics for the Real Time Online Traffic Classification. *Journal of China Universities of Posts and Telecommunications*. 18(3): 92–98.

[18] Chen, Z. X., Yang, B., Chen, Y. H., Abraham, A., Grosan, C., Peng, L. Z. 2009. Online hybrid traffic classifier for Peer-to-Peer Systems based on Network Processors. *Appl Soft Comput*. 9(2): 685-694. doi:DOI 10.1016/j.asoc.2008.09.010.

[19] Che, X. H., Ip, B. 2012. Packet-level Traffic Analysis of Online Games from the Genre Characteristics Perspective. *J Netw Comput Appl*. 35(1): 240–252. doi:DOI 10.1016/j.jnca.2011.08.005.

[20] A, H. M. 1998. Correlation-based Feature Selection for Machine Learning. Waikato University.

[21] Chen, Y., Yang, B., Dong, J. 2004. Nonlinear System Modelling Via Optimal Design of Neural Trees. *Int J Neural Syst*. 14(2): 125–137.

[22] Min, D., Xingshu, C., Jun, T. 2013. Online Internet Traffic Identification Algorithm Based on Multistage Classifier. *Communications China*. 10(2): 89–97. doi:10.1109/cc.2013.6472861

[23] Witten, I. H., Frank, E. 2005. *Data Mining Practical Machine Learning Tools and Techniques.* Diane Cerra,