

Document downloaded from:

<http://hdl.handle.net/10251/116174>

This paper must be cited as:

Canovas Solbes, A.; Jimenez, JM.; Romero Martínez, JO.; Lloret, J. (2018). Multimedia Data Flow Traffic Classification Using Intelligent Models Based on Traffic Patterns. IEEE Network. 32(6):100-107. doi:10.1109/MNET.2018.1800121



The final publication is available at

<http://doi.org/10.1109/MNET.2018.1800121>

Copyright Institute of Electrical and Electronics Engineers

Additional Information

Multimedia Data Flow Traffic Classification using Intelligent Models based on Traffic Patterns

Alejandro Canovas, Jose Miguel Jimenez, Oscar Romero, Jaime Lloret

Instituto de Investigación para la Gestión Integrada de Zonas Costeras, Universitat Politècnica de València, Spain,

alcaisol@posgrado.upv.es, jojher@dcom.upv.es, oromero@dcom.upv.es,
jlloret@dcom.upv.es.

Abstract—*Nowadays, there is high interest in modeling the type of multimedia traffic with the purpose to estimate the network resources required to guarantee the quality delivered to the user. In this work we propose a multimedia traffic classification model based on patterns that allows us to differentiate the type of traffic by using video streaming and network characteristics as input parameters. We will show that there is low correlation between network parameters and the delivered video quality. Because of this, in addition to network parameters, we also added video streaming parameters in order to improve the efficiency of our system. Finally, it should be noted that, based on the objective video quality received by the user, we have extracted traffic patterns that we will use to perform the development of the classification model.*

Keywords: *Artificial Intelligence, traffic patterns, Multimedia, QoE.*

I. Introduction

Nowadays, video streaming applications have developed significantly. For this reason, the type of traffic modeling has a great interest in order to guarantee both the Quality of Experience (QoE) and the Quality of Service (QoS). There are different metrics for the objective video quality evaluation. The Peak Signal to Noise Ratio (PSNR) parameter is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Normally, higher PSNR indicates that the reconstruction is of higher quality [1]. According to the mathematical equations for calculating Mean Square Error (MSE) and PSNR, it can be inferred that they represent similar error values, i.e., the calculated error is of the same degree. Because of this, PSNR can be considered as an unofficial representative of all the above mentioned video quality metrics, and still the most widely used metric for video quality estimation in many video processing systems [2]. The Human Visual system (HVS) is highly adapted to extract the structural information from the area of viewing. Structural Similarity Index Measure (SSIM) uses this characteristic of the HVS for quality estimation

of the processed digital video. Structural information of an image can be defined by those characteristics that represent the structure of the objects in the scene, independently of the mean brightness and contrast [2]. These measurements are based on three components: luminance comparison, contrast comparison and structure comparison [3]. New Quality Index (NQE) works like SSIM index. NQE defines picture distortion as a combination of three factors: difference in mutual characteristics, difference in luminance and difference in contrast. Human eye sensitivity to spatial-temporal pattern decreases with high spatial and temporal frequency. Based on different sensitivity, high spatial or temporal information can be represented with less data and less precision, while human eyes are more or less insensitive to the loss of this information. This characteristic of HVS is exploited by Discrete Cosine Transform (DCT) quantization, which is the base for Video Quality Metric (VQM) [1]. The values of VQM start from 0 and can reach around 12 in real situations. VQM value of 0 represents minimum distortion and maximum quality [3]. In conclusion, SSIM metric has quite better performance compared to PSNR and, in most cases, it performs very similar to the Human Visual System. But, imperfections are also present. SSIM is almost insensitive to changes in brightness, contrast and hue. And, when these changes are bigger, SSIM values can become largely inverted. VQM mostly considers the changes that are more noticeable to the human eye. We will use NQE, VQM, SSIM and PSNR metrics to conform traffic patterns.

In this paper, we propose a multimedia traffic classification model based on patterns by using video streaming and network characteristics as input parameters and video quality evaluation measures as output. It uses different learning methods to find which of them provides the best result, and therefore, better fits the problem. The methods we have employed are based on neural networks, vector support machines, statistics and the nearest neighbors. The statistical method (Kernel) is based on Discriminant analysis [4]. This term is broadly used to include problems associated with the (statistical) separation between distinct classes or groups. It includes a wide range of problems in statistical pattern recognition, where a pattern is considered as a single entity and is represented by a finite dimensional vector of the pattern features. The neuronal network (NN) method is based on adaptive networks. One of the adaptive networks class makes the interpolation scheme explicit. For this class, the property of learning is equivalent to the solution of a set of linear equations. Thus, these networks represent nonlinear relationships while having a guaranteed learning rule. Another learning method used is the Support Vector Machines (SVM). The machine implements the following idea: input vectors are nonlinearly mapped to a very high dimension feature space. In this feature space, a linear decision surface is constructed. Special properties of the decision surface ensure high generalization ability of the learning machine. Finally, we used a classification method based on distances, in particular, K Nearest Neighbor (KNN). We have used objective and subjective video quality evaluation to get traffic patterns. The subjective quality is defined by the perception of the user. This is measured in Mean Opinion Score (MOS), from 1 to 5, where 1 is perceived as very bad quality, while 5 is considered very good quality. On the other hand, the objective quality is defined when mathematical models are used to evaluate the video quality [5].

This work is distributed in the following way. Section 2 corresponds to the state of the art of the object of study. In this section we compare our work with the related work. In section 3 we make a description of the testbed we have used to extract data for further analysis. In section 4 we explain how we have carried out the development of the intelligent classification method for traffic modeling based on patterns. Finally, in section 5 we analyze the obtained results, the conclusions, and our future work.

II. Related Work

Today, as a consequence of the growing volume of voice, video and other data traffic in real time, multiple works have been proposed to implement an adequate management of the classification of the network traffic that traverses operators and companies. This big amount of traffic makes necessary traffic classification for both safety issues and service agreements, including billing and QoS.

Usually, operators rely on deep packet inspection (DPI) to identify well-known protocols. When using DPI techniques, the operators use the signatures of the applications that usually are obtained by inspecting the packets or by reverse engineering. But this solution is not totally effective. For example, due to the appearance of new applications or the change of signature when updating applications, it is not possible to maintain an updated signature database.

Some authors like in [6] and [7] propose to automatically modify the formats at their origin to detect them later. Other authors, such as in [8] propose SANTaClass, a self-adapting real-time classification system. According to the authors, SANTaClass can be connected to any network, to learn the signatures of network applications without the need for supervision.

There are authors who propose to use Machine Learning (ML) to classify Internet traffic. Normally, different methods are used for the selection of traffic characteristics. These methods are usually classified into the following three categories: filter methods, wrapper methods and hybrid methods.

Nguyen et al. in [9] carried out a study on eighteen papers presented, between 2004 and 2007, about IP traffic classification using ML. In addition, they studied the key requirements for using ML in traffic classification, and check how many requirements are used in each of the eighteen studied papers.

Moore et al in [10] apply a Naïve Bayes estimator to classify the traffic according to the applications. They classify traffic only as input from the Naïve Bayes estimator. They show that a high level of accuracy can be achieved, and with two refinements, they announce an

accuracy of 95%.

Soysal et al. in [11] compare the algorithms Bayesian Networks (BNs), Decision Trees (DTs) and Multilayer Perceptrons (MLPs), to classify traffic in six different classes: Peer-to-Peer (P2P), web (HTTP), content delivery (Akamai), bulk (FTP), service (DNS) and mail (SMTP) traffic flows on the Internet. The authors state that the BNs and DTs algorithms can be used effectively for traffic classification.

Jiang et al. in [12] propose an architecture field-programmable gate array (FPGA) based. According to the authors, the FPGA accelerates the statistical characterization of multimedia applications, and also ensures great accuracy in the classification of traffic.

Ertam et al. in [13] use Kernel Extreme Learning Machine (KELM) to classify Internet traffic. In the method developed by the authors, parameters selection is based on Genetic Algorithm (GA), and the algorithm WK-ELM is applied, which is based on the wavelet function. The authors highlight that the performance of their method improves the classic ELM and the WK-ELM.

In their work, Egea et al. [14] propose the modification of the Fast-Fast-Correlation Feature (FCBF) algorithm for the characterization of IoT traffic in industrial environments. Authors divide the space of characteristics into fragments of the same size, increasing the correlation and therefore increasing the ML applications that operate in the nodes. This allowed them to separate and prioritize the multimedia traffic of sensor data traffic, in such a way that they efficiently detect emergency situations. The authors Zhen and Qiong in [15] propose a new method for the selection of filter characteristics that they call BSF for ML. They compare BSF with FCBF and announce that it is beneficial to maintain the balance in the multi-class classification.

III. Test Bed Description

In this section we describe the test bed and procedure used in this work to analyze video streaming traffic. We will specify the way for sending traffic, parameters used to model the packet transmission over the network, and the way to capture the traffic in the receiver client. Then, the multimedia traffic received by the client will be analyzed, and classified according to its characteristics.

A. Test bed equipment

The test bed equipment consists of three computers, a gigabit Ethernet switch and a wireless access point, interconnected as shown in Figure 1. The operative system in all computers is Linux, Ubuntu version 16.04. The Server used to send video streaming had an Intel Core i7 processor and 16 GB RAM memory. Client and Shaper computers have Intel Core i5 processor and 8 GB RAM memory. We used a HP Procurve 3500, a gigabit Ethernet switch. The wireless access point was Linksys N Ultra Range Plus WRT320N. Both client computer and wireless access point have 802.11n wireless interfaces.

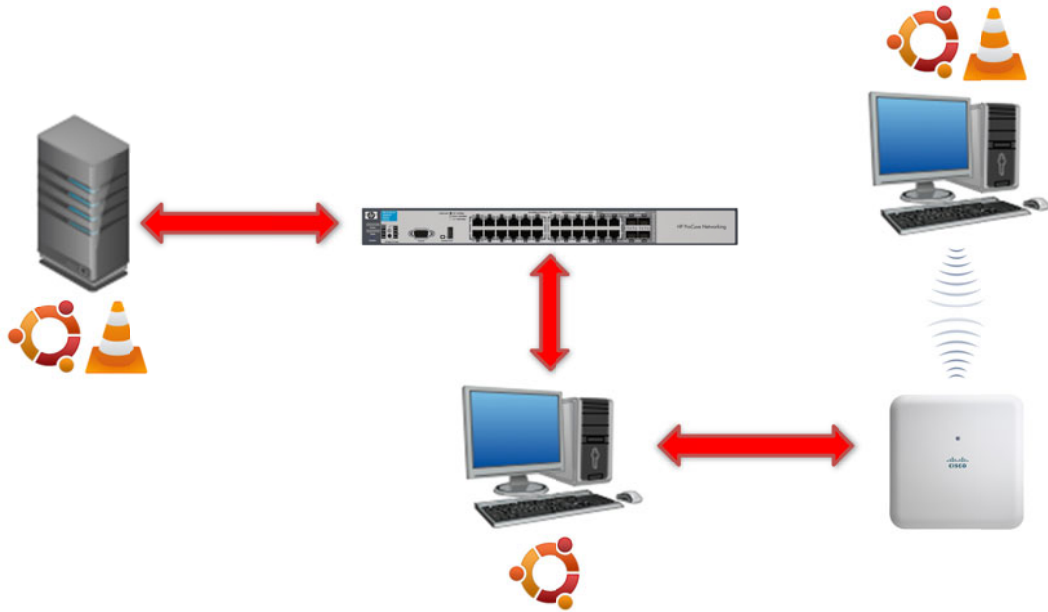


Figure 1. Test bed diagram.

The Server sends video streaming at different resolutions (R) and frame rates (F). The lowest resolution is 800x600 and the highest is 1920x1080 pixels. The rates used are 30 and 60 frames per second. When the video packets arrive to the Shaper computer, packet loss rate (L) and jitter (J) will be applied. The Shaper has been configured to emulate packet loss and jitter using the Linux traffic control tool (TC qdisc option). We have simulated packet loss from 0 to 1% using several steps. Also, the simulated jitter varies from 0 up to 60 milliseconds. Several tests have been performed by using different parameters values. First, only packet loss has been simulated. In the second set of tests, jitter has been simulated. Then, in the last set of tests, both packet loss and jitter have been simultaneously simulated, in addition to several video resolutions and frame rates.

Both Server and Client computers were running VLC software to send and receive video streaming respectively. The client was also saving the received video for further analysis. The gigabit Ethernet interfaces of the Shaper were in bridging mode. The Client was connected to the wireless access point at 150 Mbps.

B. Set of data extraction

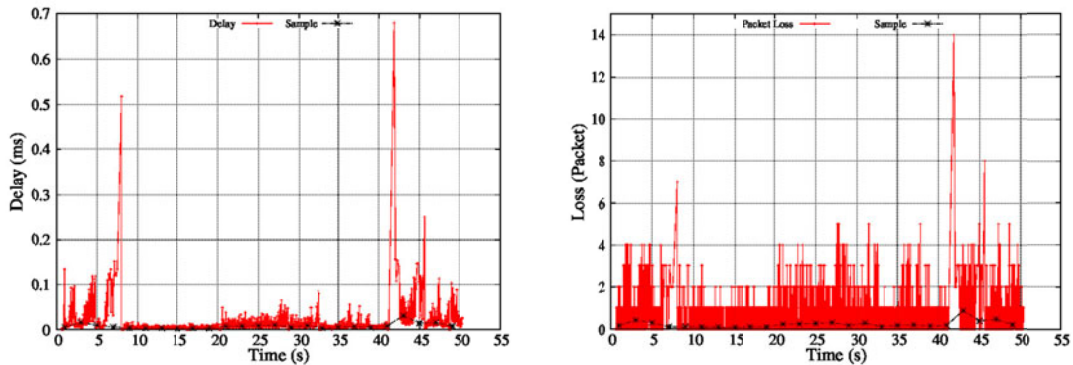
In this subsection we will explain how the extraction of the data set has been performed. We have made a total of 495 experiments. For each experiment we have recorded the received video. Therefore, for each test we have the video broadcast without degradation and the video received with degradation. Thus, by having both videos we have been able to carry out an objective study of the perceived image quality. As a result of this study we have obtained values of PSNR, NQI, VQM and SSIM. The target is to obtain an approximation of the quality perceived by the user using those obtained values. To extract the characteristics of the multimedia video stream, we have split every video into 2 seconds video clips. For each one, we have extracted information of the type: Stream size,

Proportion of this stream, Data Size, Footer Size, Bit rate, Width, Height, Framerate and Minimum frame rate. Finally, in order to obtain the network parameters, Wireshark analyzer was running in the user network. At the end of each experiment we have extracted different network parameters such as jitter, delay, packet loss and bandwidth.

IV. Classification Intelligent Model

A. Pre-processing of data

For the pre-processing of data, we perform a sampling to characterize the network status, data flow and video quality received during different multimedia transmissions. For this purpose, we first divide each video transmission into a group of pictures (GOPs) equivalent to 2 seconds video clip. For each clip, we extract the average value for network, flow and video quality parameters. In this way, for each transmission of 50 seconds we will extract 25 samples that will characterize the type of traffic based on the network status, multimedia video streaming and quality received by the user. In Figure 2 we can see the result of this process for a video transmission with a resolution of 1280x1024 at 30fps and a link bandwidth of 10Mbps.



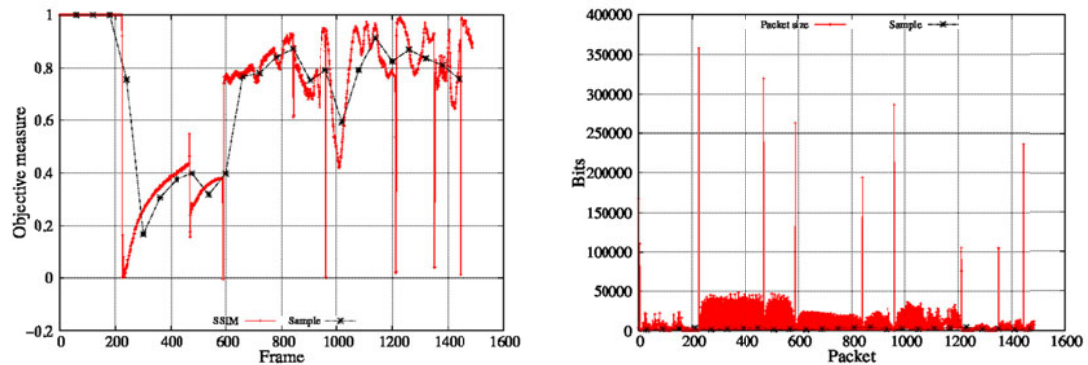


Figure 2. Samples extraction from network parameters, video quality received and video streaming during a transmission. From left to right, and from top to bottom: (a) delay sampling, (b) packet loss sampling, (c) received quality sampling based on SSIM and (d) size of received video flow.

As we can see in Figure 2 (a), there are two important delay peaks during the seconds 7 and 42 corresponding to changes of scene in the transmitted video, with values of 0.5 and 0.7ms respectively. At the same moment, packet loss peaks also occurs, as can be observed in Figure 2 (b). In this case, the peaks correspond to packet losses of 7 and 14 packets respectively. The average packet loss during the transmission for this video is usually between 1 and 3 packets per second. In Figure 2 (c), the SSIM values of the video stream received are shown. From second 7 to 20, the degradation of the video is considerable, with values lower than 0.4. This time interval coincides with the set of frames between 200 and 600. The rest of inverted peaks coincide with I-frames or scene changes. The latter can also be seen in Figure 2 (d), which shows the size of packets during transmission. If we analyze this figure, we see that generally the total packets data do not exceed 500 Kbytes, except those that have I-frames ranging from 1Mbyte to 3.5Mbytes. In all the cases that have been analyzed, after the sampling process we obtain another measure that represents the original. This can be seen easily in Figure 2 (c), where the SSIM sampling is represented.

B. Traffic patterns extraction

There are two ways to analyze the quality of video received by the user: objective and subjective. The objective method is obtained by mathematical formulas, and the subjective method, by the perception of the observer. For traffic classification model we use both methods. First, we obtain several patterns of the same video flowthrough different calculations of objective quality. Then, we determine which of these patterns are closest to the value of the MOS through subjective evaluation. For example, in the following figure we have extracted two patterns from the same video stream in the same conditions as in previous section (Figure 2).

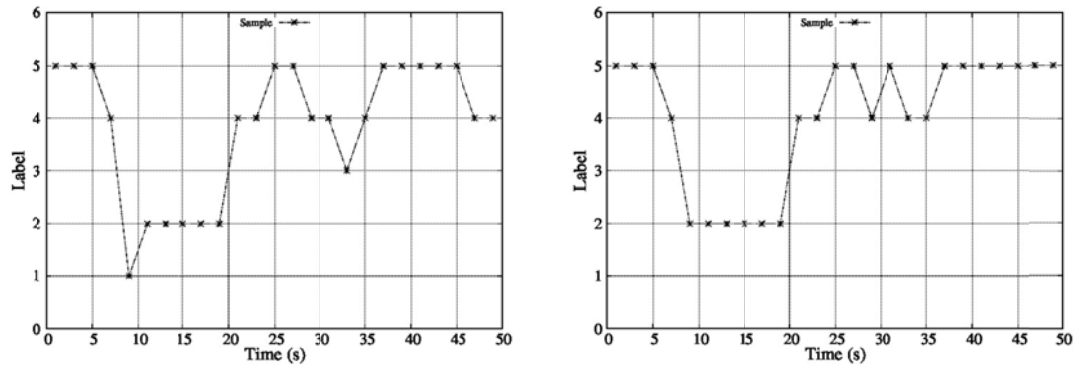


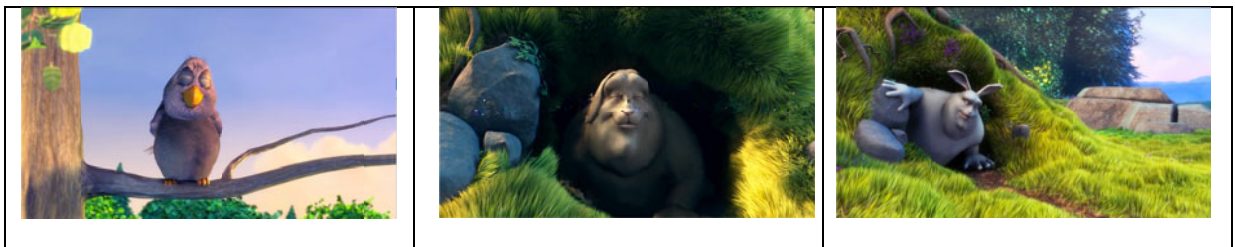
Figure 3. Multimedia traffic quality patterns extraction. Left (a), pattern based on ISSM. Right (b), pattern based on PSNR, ISSM, VQM and NQI combination.



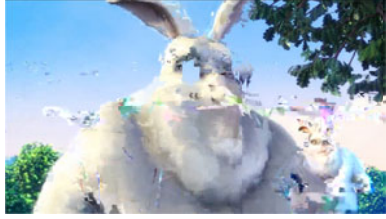
Figure 3 (a) shows the traffic quality pattern based on ISSM. It can be observed that from second 7, quality of the video decreases considerably. We have considered the quality in MOS values, that is, 5 means very good quality and 1 means very bad quality. We can also observe that during second 33 approximately, the image quality decreases again. However, when we analyze the quality of the video in a subjective way we do not observe any degradation. The result obtained in Figure 3 (b) coincides with the observation that we have just considered. This analysis process has been repeated with different video streams to finally reach the conclusion that, the extraction of patterns based on the combination of PSNR, ISSM, VQM and NQI is the one with the highest correlation with the subjective measure of video quality. Therefore, we select this method for pattern extraction.

On next situation, we will study two important features compared with previous analysis:

- how a low value affects the image quality?
- which value we will consider to identify that the critical traffic is starting?

In figure 4 we can appreciate both aspects.



4s - OK	32s – min. Line Pixel Error	36s - OK
		
40s – Ghost+Blur	46s – min. Pixel + min. Error Chrominance.	48s - Ghost+Blur

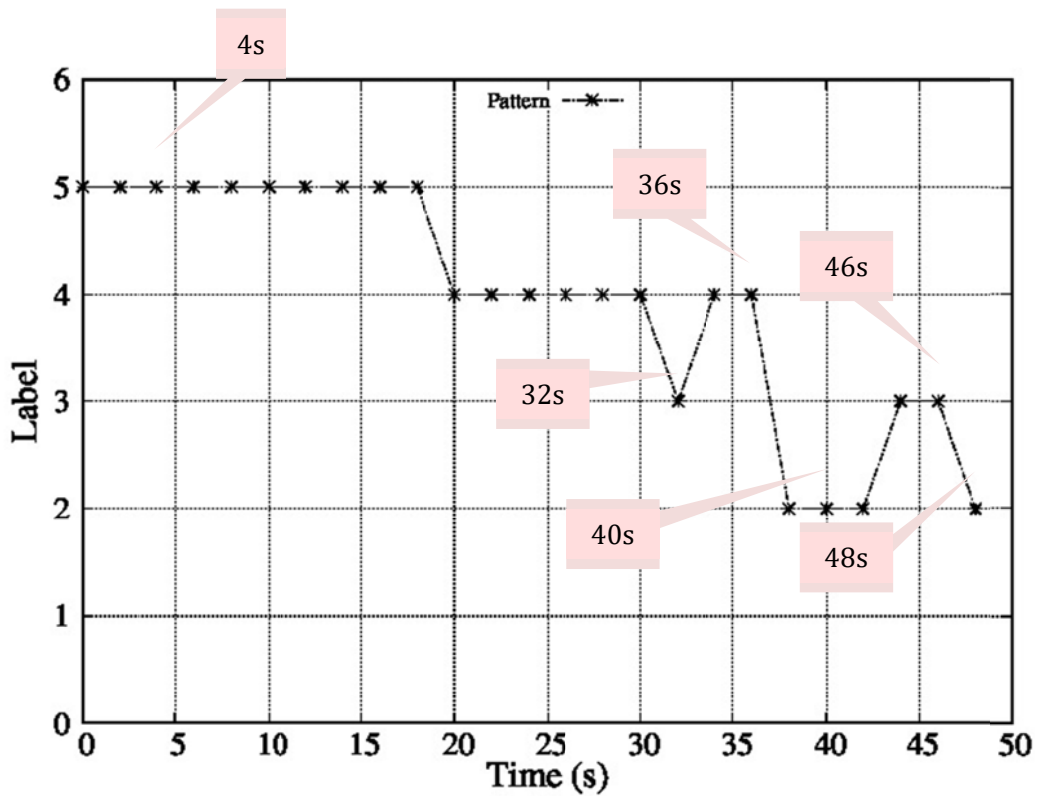


Figure 4. Video streaming objective and subjective analysis. Top, (a), video capture key moments. Bottom, (b), traffic pattern pointing key moments.

As we can see in the traffic pattern given in Figure 4 (b) from the beginning until second 18, the video is being received with the highest quality. Then, from second 18 to 30 there is a small degradation, not too much considerable. Figure 4 (a) shows a capture made in

second 4, representative for this period. Later, in second 32, as can be seen in Figure 4 (b), there is a decrease of the quality to value 3. This value has been taken into consideration as we can see in Figure 4 (a). At this moment, there is a pixelated band in the video that would affect the QoE perceived by the user. From this second, the quality of the video improves, but, in the second 38 again, the quality falls down to value 2 as we can see in Figure 4 (b). The degradation with these values is certainly important, causing errors like ghost and blur, shown in Figure 4 (a). After this degradation, quality increases to value 3, but as we can observe in Figure 4 (a), during the second 46, there are still some errors, although minimal, of pixelation and chrominance. Finally, the quality of the video degrades again during the final moments as we can see in Figure 4 (b). Quality decreases again to values 2 and, consequently, new errors of ghost and blur appear as we can see in Figure 4 (a).

After extracting traffic patterns, we study now how to extract a model that determines the type of traffic based on these patterns, by using network characteristics and multimedia video streaming as input parameters. For this purpose, we first evaluate which parameters should be selected. In order perform this evaluation, we analyze the correlation between these parameters and the labels of the pattern where they belong to. The results obtained are shown in table I.

TABLE I
CORRELATION BETWEEN NETWORK PARAMETERS AND MULTIMEDIA AND TRAFFIC PATTERNS

Coefficients	Delay	Jitter	Packet Loss	BW	Frame count	Stream size
Correlation	0.174	0.112	0.932	0.155	0.0132	-0.217
determination	0.030	0.012	1.082	0.024	0.0001	0.047

Coefficients	Proportion of stream	Footer size	Resolution	Bit rate	Framerate	Min. framerate
Correlation	-0.1194	-0.0051	-0.30	-0.17	0.148	0.475
determination	0.014	0.00002	0.094	0.030	0.022	0.226

In Table I, the level of correlation between the characteristics of the received video and the network with respect to the objective quality of the received video is shown. The measure that shows a maximum correlation is “Min. framerate”. On the other hand, “FooterSize” shows a low correlation, meaning that we will discard it in the next study, which consists of extracting the multimedia traffic classification model.

C. Traffic classification model extraction based on patterns

After pre-processing the samples, we have a total of 2741 samples with 14 characteristics per sample. From these samples we develop the classification model. For this target, we first perform a learning process and then, a test process to validate the model. Learning supervision is based on traffic patterns. There are 5 labels that represent the type of traffic during the video transmission:

- Non critical
- Low critical
- Some critical
- Critical
- Very critical

As shown in previous section, the extraction of these values is extracted based on the objective QoE. The range of these values is from 1 to 5, where 1 means very bad perceived QoE, and corresponds to a very critical label. Value 5 means very good quality traffic, and corresponds to a non critical label. We develop a classification model that will include QoS values and video streaming characteristics as well as QoE objective discrete values.

Once we have the labeled and pre-processed samples, we can initialize the learning process. For this procedure, we will use 80% of the samples for training, and the remaining 20% for testing. The selection of each group, training and test, has been randomly selected from the all samples. During the learning process, the stop criterion used is based on the mean square error and the number of cycles.

In order to extract the model, we have analyzed several learning methods to find out which of them provides the best result. The methods we have used are based on neural networks, vector support machines, statistics and the nearest neighbors. After the learning process with each proposed methods, we will evaluate the obtained results. For this, we use measure Precision (P), Recall (R), Accuracy (A) and F1, being F1 the harmonic mean of precision and recall. With these measures, we determine the accuracy of our system, based on the following classification results:

- TP – The model classifies the traffic as critical and it matches.
- TN – The model classifies the traffic as non critical and it matches.
- FP – The model classifies the traffic as critical but it does not match.
- FN – The model classifies the traffic as non critical, but it is critical.

Where, **T**= True, **F**= False, **P**= Positive, **N**= Negative and considering *critical* as Some critical/Critical/Very critical traffic and *non critical* as Non critical/Low critical traffic. Figure 5 shows the results obtained.

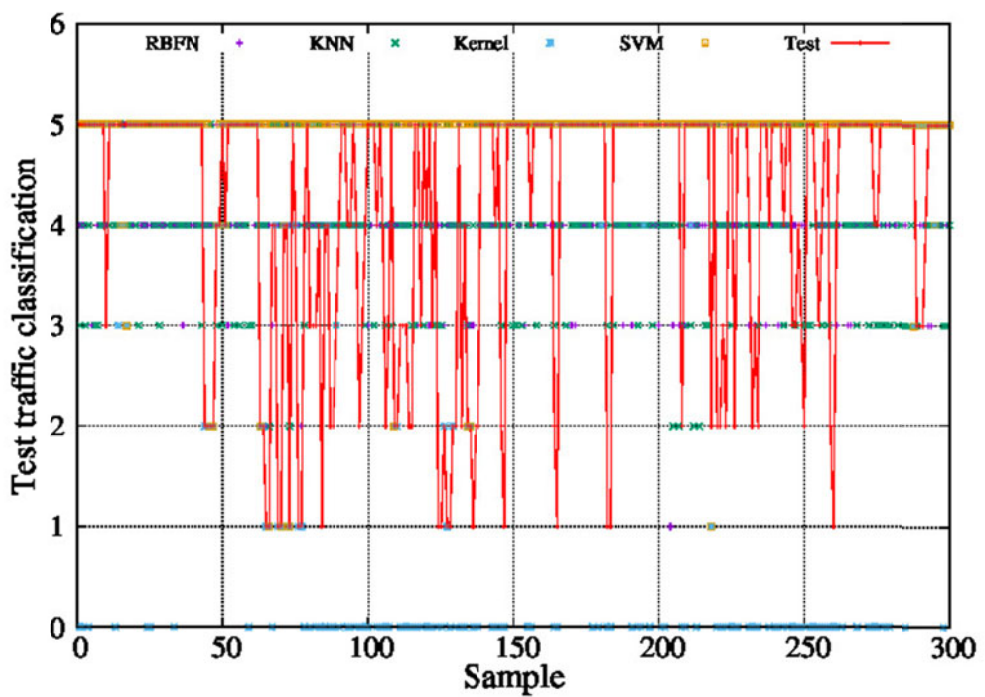
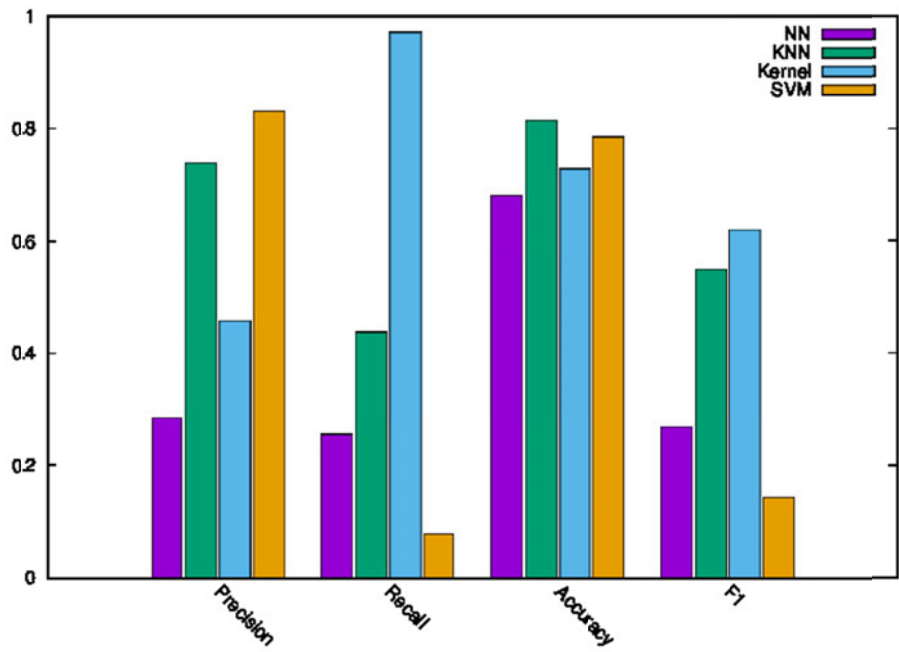


Figure 5. Results of classification based on the different methods. Top, (a), Accuracy

level. Bottom, (b), Traffic Classification.

As we can see in Figure 5, the KNN-based model has the best results, with 81% of accuracy level. This measure is based on the total set of observations. However, for the development of our system, it is of great interest to observe which method is more precise by placing special emphasis on false negatives, that is, when the model classifies traffic as non-critical but is not true. The measure that most takes this aspect into account is the Recall. For this case, the results obtained with the KNN and neural networks provide a value of 43% and 25% respectively. For the SVM-based method, results are clearly worse, with a value of only 7%. On the other hand, the method that gives very good results is the statistical method, obtaining a value of 97%. These last two cases are of great interest. Figure 5 (b) shows that the statistical method has more precision, with low values, but as we can see, the tendency is to select low values. Thus, for metrics that take more into account FP, such as accuracy, the number of matches will be considerably reduced. The opposite will happen with SVM. In this case, although it does not get very good result for critical traffic, it is very accurate. As shown in Figure 5 (a), for F1, the result obtained between KNN and Kernel is quite similar. With KNN we obtained 54% and with Kernel, 62%. The results were worse for NN, with 26%, and for SVM on only 14%. With NN, very similar results have been extracted in terms of accuracy, Recall and F1. In precision, the value obtained was 28%. This result is slightly improved with the statistical method, with 45%, and considerably more with SVM and KNN methods, with 83% and 73% respectively. Regarding to accuracy, unlike with Recall, the measure is analyzed with respect to FP. This aspect would make it less determinant for a traffic classification system in which it was very important to determine if traffic is critical, since the important issue is not to get a wrong result when it is a critical factor. It can be said that F1 is taken into account for both FN and FP. Finally, as shown in Figure 5 (a), the KNN-based method could be considered more reliable in general computing, but analyzing Figure 5 (a) and (b), SVM is more accurate because it classifies critical traffic samples while is not possible with KNN.

V. Conclusion

In this paper, we have proposed a multimedia traffic classification system based on patterns. First, we have extracted patterns used to model the objective quality perceived by the user. Then, based on these patterns, we have developed a classification system to classify the quality of the traffic. The input parameters for this system are network parameters and video streaming characteristics.

One of the main problems of this study has been the data synchronization, since measures have been taken at frame and time level. In reception, times vary with respect to the transmission times. This problem was solved using different scripts.

Once the patterns have been extracted, they have been used to develop the intelligent classification system. Based on the results obtained, the SVM classification is more

accurate than with the other methods. We have analyzed other measures such as Recall, F1 and precision. The measure of Recall has special interest because it takes into account false positives more than the other measurements. With this measure, the classification method that has improved the results is the statistical method or kernel. Generally, good accuracy results have been obtained from all methods. On the other hand, this is not the case of the rest of the measurements. It is noticeable that that when false negatives are taken into account, the results get much worse for SVM. And when false positives are taken into account, results get worst for kernel. Neural networks only provide good results if the whole set of observations are taken into account.

Finally, in future work we will use different sampling frequencies based on the GOP, with different type of videos according to their characteristics, formats, encodings but also different videos to the original one. Moreover, we will try to improve the results of both, accuracy and recall. For example, to improve the accuracy in SVM we will analyze other parameters such as the activation function, gamma, C, etc. On the other hand, we think that adding more cases with critical traffic, recall could be improved, since cases of FN would occur in an easier way. Finally, we will make a further study of traffic patterns, with the purpose of analyzing the interrelation between these patterns.

Acknowledgment

This work has been supported by the Ministerio de Economía y Competitividad in the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento within the Project with reference TIN2017-84802-C2-1-P.

References

- [1] P. Symes, Digital Video Compression, McGraw-Hill, New York, 2004.
- [2] Z. Wang, A.C. Bovik, H. R. Sheikh, E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, April, 2004.
- [3] Kotevski, Z & Mitrevski, Pece. Performance Assessment of Metrics for Video Quality Estimation. Conference on Information, Communication and Energy Systems and Technologies (ICEST'2010), June 23-26, 2010. Ohrid, Macedonia, Pp. 693-696.
- [5] McLachan, G. J. Discriminant Analysis and Statistical Pattern Recognition, Wiley. Indianapolis, USA, 1992.
- [4] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "Study of Subjective and Objective Quality Assessment of Video," in IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1427-1441, June 2010. doi: 10.1109/TIP.2010.2042111
- [6] Caballero J., Yin H., Liang Z., Song D., Polyglot: automatic extraction of protocol

message format using dynamic binary analysis, in: ACM Conference on Computer and Communications Security, Alexandria, VA, USA, October 29 - November 02, 2007, <http://doi.org/10.1145/1315245.1315286>

[7] Wang Y., Yun X., Shafiq M.Z., Wang L., Liu A., Zhang Z., Yao D., Zhang Y., Guo L., A semantics aware approach to automated reverse engineering unknown protocols, 2012 20th IEEE International Conference on Network Protocols (ICNP), 30 Oct.-2 Nov. 2012, Austin, TX, USA, <https://doi.org/10.1109/ICNP.2012.6459963>

[8] Tongaonkar A., Torres R., Iliofotou M., Keralapura R., Nucci A., Towards self-adaptive network traffic classification, Computer Communications, Vol. 56, Issue 1, February 2015, Pp. 35-46, <https://doi.org/10.1016/j.comcom.2014.03.026>

[9] Nguyen T. T. T., Armitage G., A survey of techniques for internet traffic classification using machine learning, IEEE Communications Surveys & Tutorials, Vol. 10, Issue: 4, Fourth Quarter 2008, Pp. 56 - 76, <https://doi.org/10.1109/SURV.2008.080406>

[10] Moore A. W., Zuev D., Internet traffic classification using Bayesian analysis techniques, 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems (SIGMETRICS '05), Banff, Alberta, Canada, June 06 - 10, 2005, Pp. 50 - 60, <https://doi.org/10.1145/1071690.1064220>

[11] Soysal M., Schmidt E. G., Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison, Performance Evaluation, Vol. 67, Issue 6, June 2010, Pp. 451-467, <https://doi.org/10.1016/j.peva.2010.01.001>

[12] Jiang W., Gokhale M., Real-Time Classification of Multimedia Traffic using FPGA, 2010 International Conference on Field Programmable Logic and Applications (FPL), 31 Aug.-2 Sept., 2010, Milano, Italy, <https://doi.org/10.1109/FPL.2010.22>

[13] Ertam, F., Avci E., A new approach for internet traffic classification: GA-WK-ELM, Journal Measurement, Vol. 95, January 2017, Pp. 135-142, <https://doi.org/10.1016/j.measurement.2016.10.001>

[14] Egea S., Rego A., Carro B., Sánchez-Esguevillas A., Lloret J., Intelligent IoT Traffic Classification Using Novel Search Strategy for Fast Based-Correlation Feature Selection in Industrial Environments, IEEE Internet of Things Journal, Volume: PP, Issue: 99, Pp. 1 - 10, <https://doi.org/10.1109/JIOT.2017.2787959>

[15] Zhen L., Qiong L., New Feature Selection Method for Internet Traffic Classification Using ML, Physics Procedia, Vol. 33, 2012, Pp. 1338-1345, <https://doi.org/10.1016/j.phpro.2012.05.220>

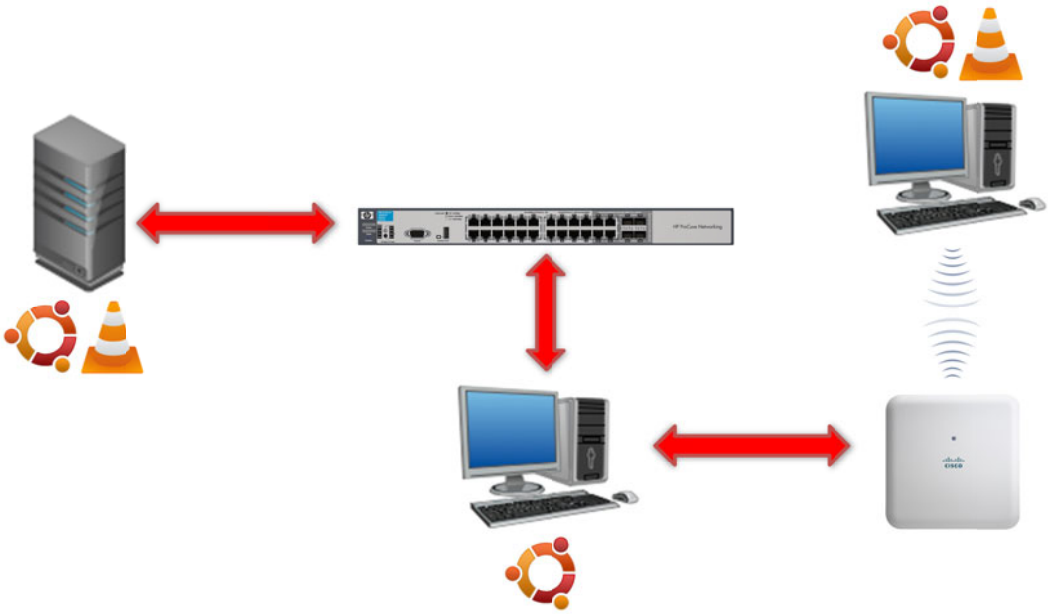


Figure 1. Test bed diagram.

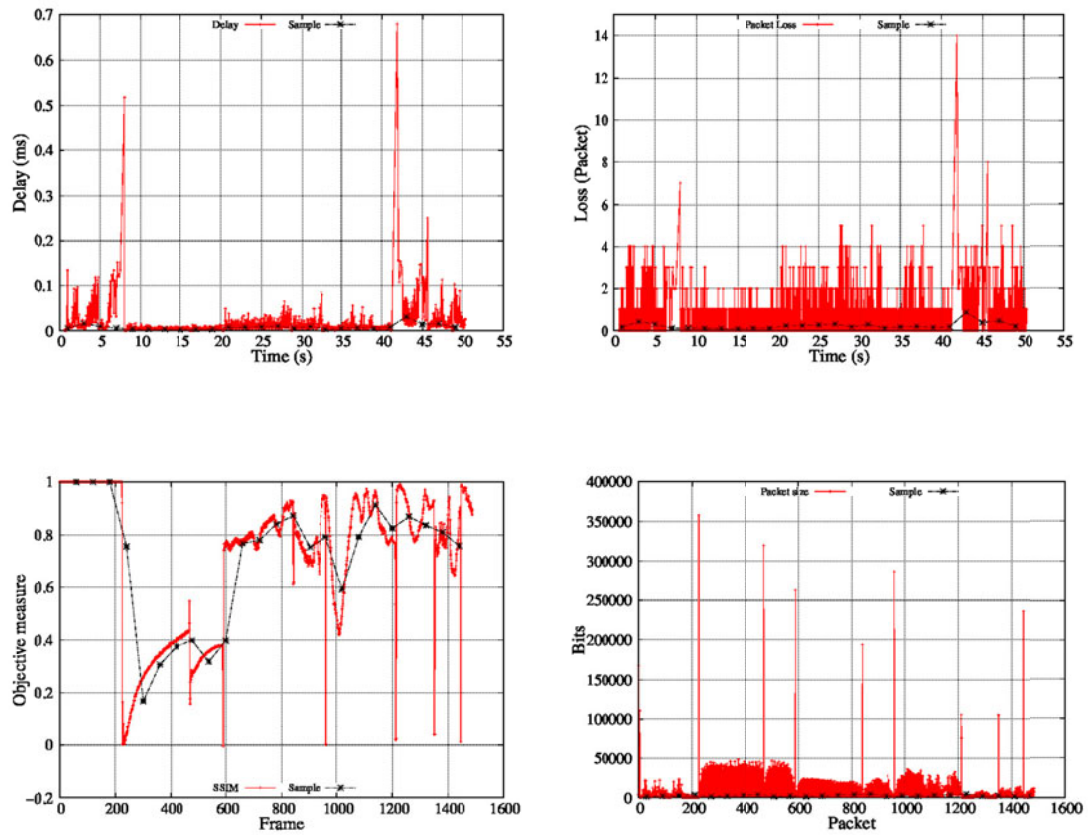


Figure 2. Samples extraction from network parameters, video quality received and video streaming during a transmission. From left to right, and from top to bottom: (a) delay sampling, (b) packet loss sampling, (c) received quality sampling based on SSIM and (d) size of received video flow.

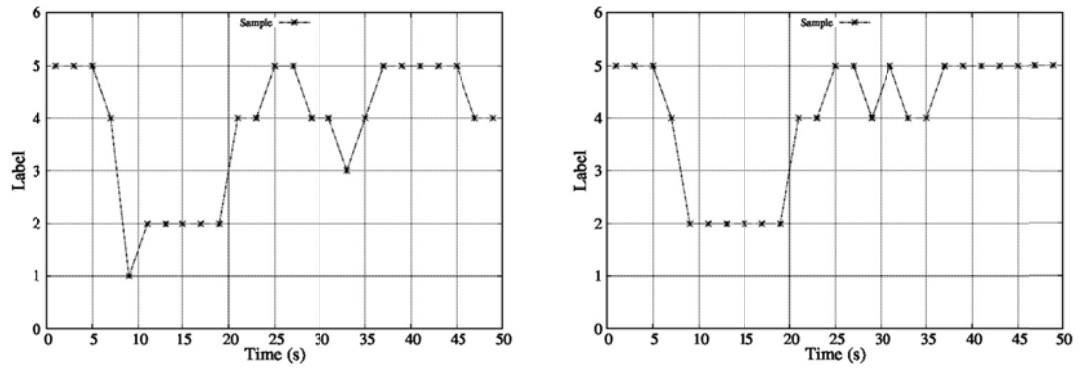
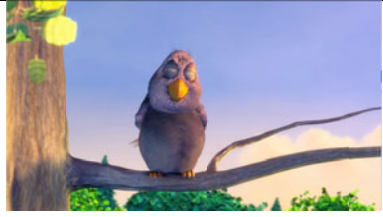

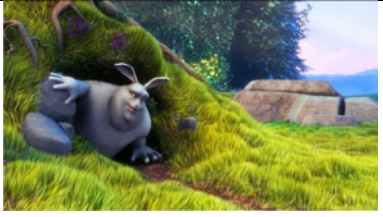


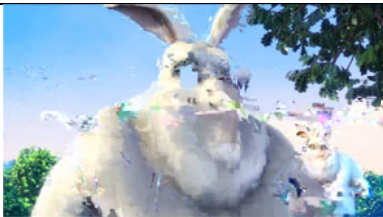


Figure 3. Multimedia traffic quality patterns extraction. Left (a), pattern based on ISSM. Right (b), pattern based on PSNR, ISSM, VQM and NQI combination.

		
4s - OK	32s – min. Line Pixel Error	36s - OK
		
40s – Ghost+Blur	46s – min. Pixel + min. Error Chrominance.	48s - Ghost+Blur

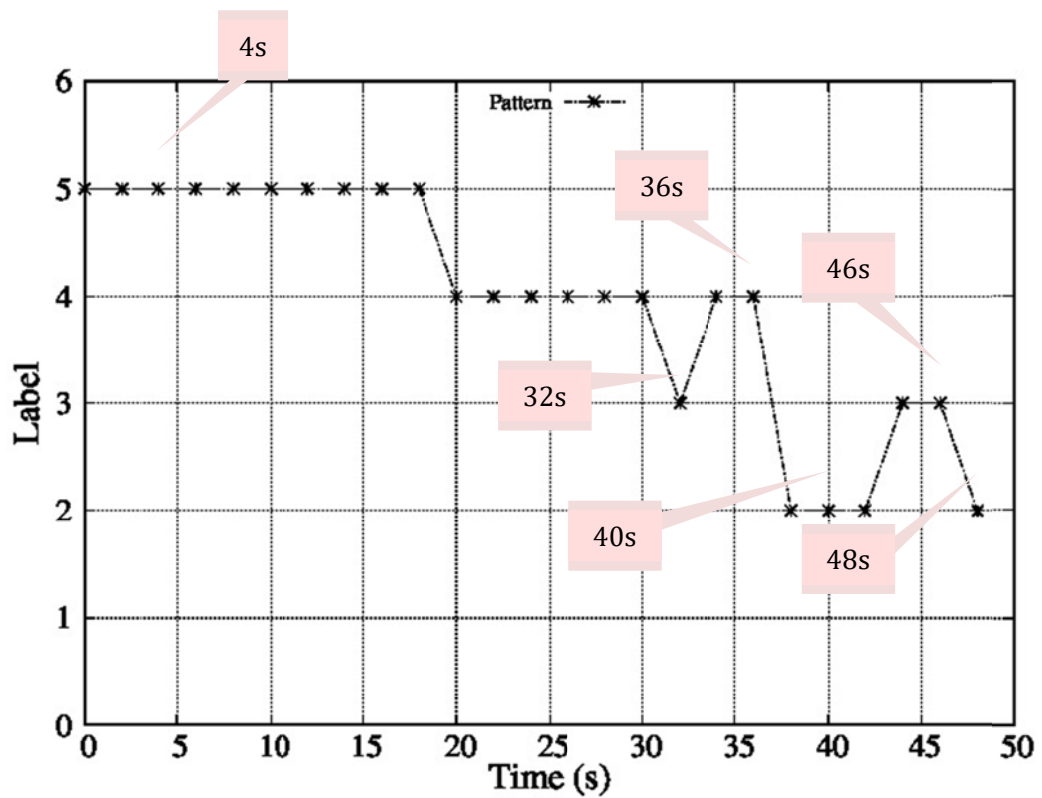


Figure 4. Video streaming objective and subjective analysis. Top, (a), video capture key moments. Bottom, (b), traffic pattern pointing key moments.

TABLE I
CORRELATION BETWEEN NETWORK PARAMETERS AND MULTIMEDIA AND TRAFFIC PATTERNS

Coefficients	Delay	Jitter	Packet Loss	BW	Frame count	Stream size
Correlation	0.174	0.112	0.932	0.155	0.0132	-0.217
determination	0.030	0.012	1.082	0.024	0.0001	0.047

Coefficients	Proportion of stream	Footer size	Resolution	Bit rate	Framerate	Min. framerate
Correlation	-0.1194	-0.0051	-0.30	-0.17	0.148	0.475
determination	0.014	0.00002	0.094	0.030	0.022	0.226

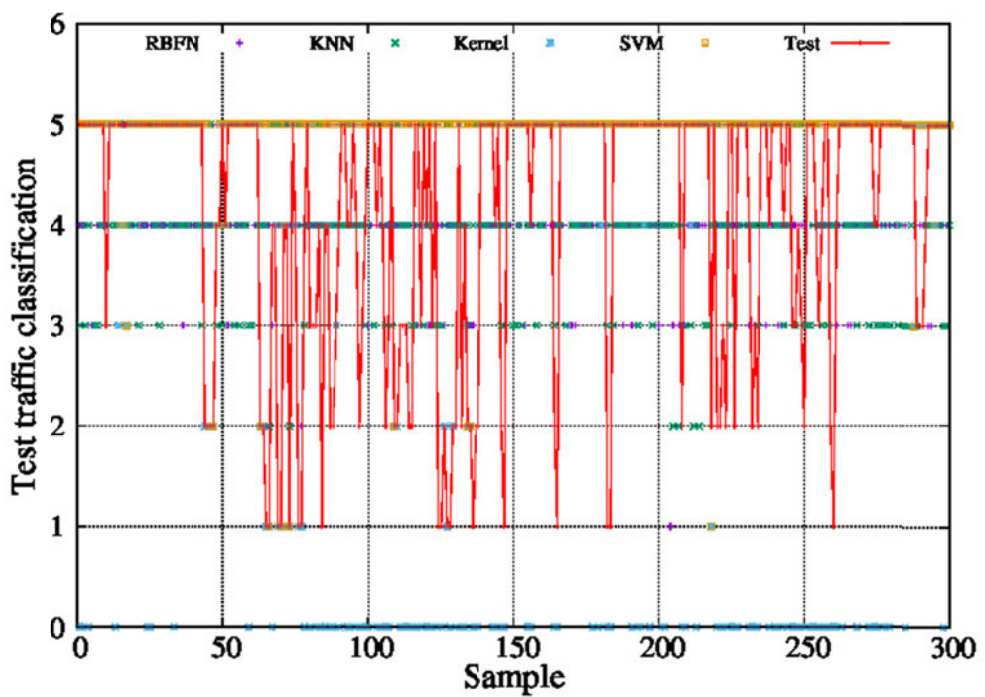
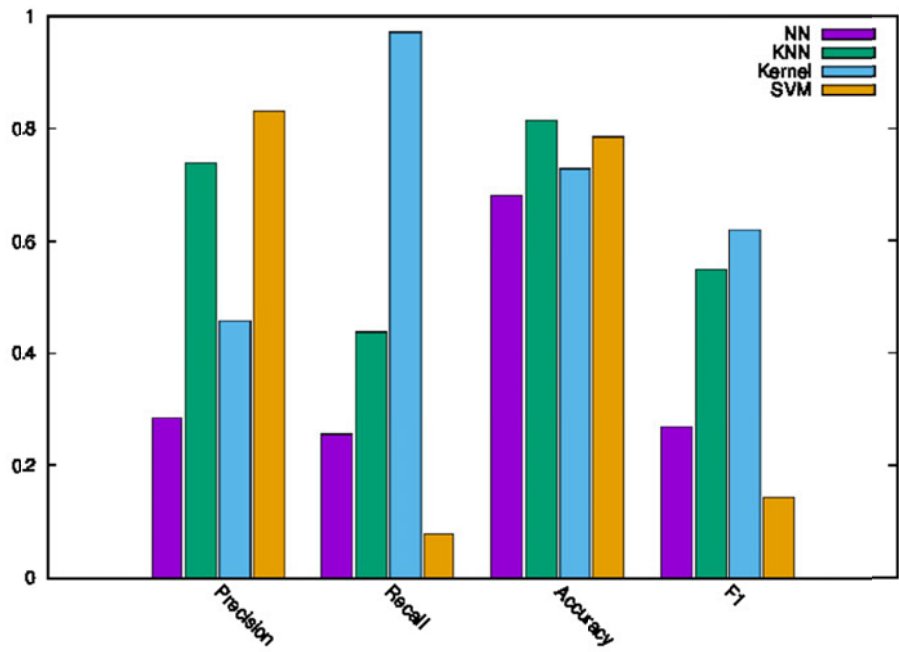


Figure 6. Results of classification based on the different methods. Top, (a), Accuracy

level. Bottom, (b), Traffic Classification.