A Multi-Classification Approach for the Detection and Identification of eHealth Applications

Monika Grajzer, Michał Koziuk, Piotr Szczechowiak
Telcordia Poland
Applied Research Center
Email: {mgrajzer, mkoziuk, pszczech}@telcordia.com

Antonio Pescapé
Department of Computer Engineering and Systems
Universitá di Napoli Federico II
Email: {pescape}@unina.it

Abstract—eHealth services category has a diversified set of traffic patterns and demands in terms of QoS assurances. Existing QoS solutions were designed to support only aggregated classes of service and cannot differentiate traffic based on an application's behavioral pattern. In order to improve the performance of eHealth applications for home and mobile users there is a need to develop new traffic identification techniques, which would work at the edge of the network. This paper addresses the above problem by proposing machine learning-based approach for eHealth traffic identification. We investigate different techniques which combine the results from multiple machine learning classifiers and show which combination of techniques is best suited for identifying diverse eHealth traffic. Our approach is validated in a mobile e-health application context and the results prove that multi-classification techniques can be used in practice to provide application-based service differentiation.

I. Introduction

Current telecommunication networks offer sufficient capacity and connection speeds to enable innovative eHealth services to both home and mobile users. eHealth and mHealth applications are gaining on popularity and network mechanisms must provide appropriate Quality of Service (QoS) assurances to support such applications. This task is challenging due to the fact that eHealth application category is very diverse and spans from simple consulting services to life-critical telemedicine applications.

The solution to the problem of providing appropriate QoS for a specific group of applications can be designed in two completely different ways. One approach is to relay only on network level mechanisms to detect applications and secure sufficient network resources for service delivery. In this case the application is not aware that the generated packets are treated in a special way inside the network. This approach requires installation of specialized traffic analysis software on edge nodes.

The second approach is based on designing application layer mechanisms, which would explicitly set the transmission priority of a given application. This approach is much harder to be implemented in reality due to the sheer number of currently available applications and software development standards. Such an approach raises also a concern in respect to the priority assignment process. It is obvious that each developer will opt for the highest priority in order to achieve the best possible performance for a given network application. Based on the reasoning presented above we decided to follow the first approach in our research work.

In current networks the traffic generated by eHealth applications is treated in the same way as in case of other applications, which are usually not that critical and far less demanding in terms of QoS. In order to implement application-based service differentiation there is a need to develop efficient traffic classification methods as well as appropriate scheduling and resource reservation techniques. In the literature we can find papers, which investigate priority assignment and scheduling techniques for eHealth applications [1]. However, such mechanisms cannot distinguish medical applications and it is assumed that eHealth traffic can be easily detected. Such an approach may lead to incorrect classification of traffic and decreased QoS for medical applications. Therefore, in this paper we focus our efforts on designing advanced traffic detection and identification techniques for eHealth applications differentiation. We also study identification of single applications, which support multiple service channels and generate diversified traffic patterns consisting of signaling packets as well as voice, video and data flows. So far this problem have not been properly addressed in the literature.

II. Related work

Traffic classification is an important part of Internet traffic engineering and has applications in several fields such as network monitoring, application identification, anomaly detection, accounting, advertising and service differentiation. Traffic classification mechanisms associate traffic flows with specific application types. There are also traffic identification mechanisms, which detect traffic generated by a specific application. Both mechanisms are closely related because classification cannot be performed without proper identification of applications. Traffic detection and classification have gained on importance in recent years due to increased interest in service differentiation and growing incentives to disguise certain applications [2].

The most basic traffic classification methods use the port numbers information included in the packet header. This classification method is simple and fast but can be very inaccurate, because current applications can hide traffic behind well known ports (e.g. using TCP port 80) or use dynamically allocated ports. More advanced methods analyze packet

payload looking for certain bit patterns. Such methods can be implemented in software or hardware and are often denoted as Deep Packet Inspection (DPI) [3], [4]. The main drawback of such methods is high computational complexity and inability to work with encrypted traffic.

The basic classification methods described above often serve as a reference (ground-truth) for assessing the performance of more advanced methods from the pattern recognition field. Most of such methods are based on Machine Learning (ML) classifiers, which make decisions based on the observation of traffic flow features [5]. ML-based techniques consider both supervised and unsupervised learning approaches. The first group of methods requires a training phase and after that assigns applications to user-defined classes. In unsupervised learning different traffic flows are grouped together to form clusters that share common features. The accuracy of ML-based techniques depends on the type of machine learning algorithm (decision tree, neural network, Bayesian analysis, etc.) as well as the set of features used for classification. High accuracy of these methods can be achieved through proper selection of training data and experimentation with the features set.

Another group of traffic classification methods look at statistical analysis [6] or at connection patterns of network flows. Such techniques are very interesting, because they operate only on the flow level and require less detailed input than packet level techniques. They can be used to identify specific applications (e.g. peer-to-peer traffic in [7]) through connection patterns. One of the main problems with such methods is non real-time operation, due to large amounts of data flows analysis. This drawback limits the applicability of such techniques for the purpose of eHealth traffic identification.

III. EHEALTH TRAFFIC IDENTIFICATION

Current networks enable all kinds of different eHealth services with new applications appearing on the market almost every day. Each of these applications have different QoS requirements when it comes to bandwidth, delay and reliability. Medical teleconsultation services require small packet delays and high bandwidth (speech and HD video streaming). Transmission of medical images (MRI, X-ray, USG) is delay tolerant but is especially susceptible to packet loss, since distorted and modified images may lead to a wrong diagnosis. On the other hand medical sensory data is not delay tolerant and even minimal packet loss is unacceptable (e.g. in case of heart rate monitoring sensors). Each of the above applications generates traffic flows, which have completely different parameters and characteristic.

Before an appropriate eHealth traffic identification technique can be designed there is a need to analyze the characteristics and traffic patterns of different eHealth applications. Based on this analysis we have defined a set of requirements, which should be fulfilled by the traffic classification mechanism:

- Real time operation – eHealth traffic is related to time-critical applications, where delays should be kept to minimum. this feature enforces real-time operation of traffic classification algorithms.
- Low number of necessary packets – a decision regarding classification needs to be made on the shortest possible part of the flow (early detection).
- High precision – correct identification must be performed with high probability and low false-positives ratio.
- Ability to identify single applications with multiple service channels – this is a characteristic feature of many eHealth applications.
- Small processing overhead – a lightweight solution is required, but higher overhead is acceptable in the offline phase.
- Ability to classify currently unknown application types – since new eHealth applications are constantly emerging, the classification method should also have the ability to recognize previously unknown applications.
- Ability to work with encrypted traffic – eHealth applications require a secure channel and apply encryption while transmitting private data.

The above set of requirements disqualifies certain traffic classification techniques. Port based methods are fast but not accurate enough to distinguish particular applications (many applications use random ports). DPI methods cannot be used, due to encrypted traffic analysis and privacy concerns. Machine learning based approaches seems to be the most promising way for eHealth traffic identification. These methods take into consideration only statistical properties of a flow, such as average packet size, packet inter-arrival time, etc. ML methods apply complex classification techniques, which make decisions based on multi-criteria reasoning without looking deep into the packet content. In this way the privacy of users data is intact. Additionally ML-based techniques can operate close to real-time [8] and can achieve high accuracy comparable to DPI methods [5]. They can also classify new applications thanks to the learning capabilities of unsupervised techniques enhanced with a training phase allowing for proper labeling.

Based on the study presented in [9] it can be observed that the accuracy of different ML classifiers varies depending on the type of analyzed traffic. Since different techniques perform better on some traffic classes it is potentially beneficial to combine standalone classifiers in a multi-classification system [8]. Such a system uses intelligent combining techniques, which learn from historical behaviors of individual classifiers. It can achieve higher accuracy than any single classifier and is more robust to changes in the mix of classified applications. In general the process of combining classifiers increases the computational complexity of the classification mechanism. On the other hand it can reduce the amount of traffic required for accurate classification (e.g. using two packets per flow instead of four) resulting in faster classification. The above features makes such mechanisms particularly suitable for identification and classification of eHealth traffic.

IV. METHODOLOGY AND TOOLS USED

The first step in the traffic classification research is to obtain the traffic traces being the subject of classification. For this purpose traffic generated by a real-world eHealth application was examined. Due to the large number of possible eHealth applications in this study, we focused only on the mobile eHealth use case. The prototype eHealth application (see Fig. 1) was developed for the Android OS and implemented three modes of communication - medical files transfer, sensor data transmission and video teleconferencing. In the first case typical MRI images of approx. 1 MB were used as the source of data. Sensor application generated data streams of approx. 120 Kbps, which correspond to data rates of real medical sensors reported in [10]. Video-conferencing used the H.264 and Speex codecs with peek data rates of 500 Kbps similar to the parameters of a real telemedicine application used in [11]. These three modes of operation are good representatives of real mobile telemedicine applications and reflect the traffic diversity of different eHealth services. Each mode was implemented as follows:

- Video-conferencing feature was based on the open source Linphone Android client enhanced with features allowing for additional communication between the users. SIP proxy/registrar server by Tivi was used to enable the signaling functionality.
- The transmission of sensor data was implemented using UDP datagrams. Each sensor sample was packed into a separate packet. The communication between mobile devices relied on an intermediate proxy used for user registration. The sensors were emulated using text files with appropriate medical data from [12].
- File transfer was implemented using TCP sockets. The communication also made use of the intermediate proxy server. The files were selected as medical images from the PhysioNet database [13].

eHealth traffic classification requires also background traffic traces which reflect the structure of regular network traffic. Moreover, the traffic traces have to be coupled with the ground truth information which enables validation of classifier performance. This implies that not all publicly available traces (e.g. NLANR, etc. [14]) are suitable for traffic classification, since in many cases they are anonymized. The ground-truth is usually obtained with the payload inspection methods, e.g. the so called l7 filter [15]. These methods, however, are not able to provide good results in the presence of anonymization, since anonymized packet content does not match the predefined application patterns.

For our investigations we have used our own background traffic traces generated in a real network environment. These traces reflect the results of recent studies on the Internet traffic structure [16]. The amount of eHealth traffic was intentionally kept low in comparison to other traffic sources, which is in line with the most probable real-world scenario. The trace had 2 million packets, which corresponds to 2GB capacity and 45 thousand sessions identified as bidirectional flows (*biflows*).
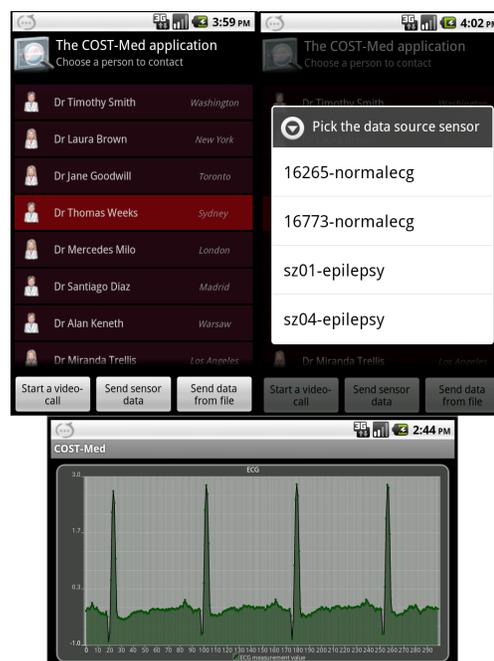

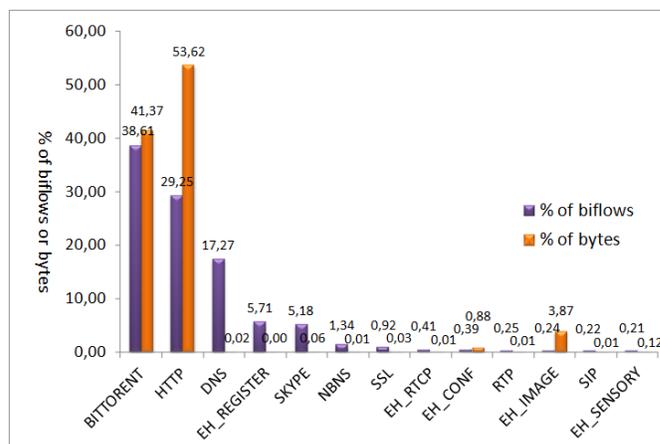Fig. 1. Prototype eHealth application running on the Android platform

Fig. 2. Traffic trace breakdown into application classes.

During the experimentation phase the l7 filter was used to provide the ground truth. All biflows identified by the filter as unknown and those with low occurrence were removed from the original trace. Moreover some sessions were identified by the filter as belonging to applications which were not present in the original trace and thus were also removed. As a general remark we tend to observe that the amount of traffic identified as P2P traffic is overrated, which suggests that also the methods widely considered as a reference do not give 100% accuracy.

We have defined 4 classes, which describe the new types of e-Health traffic that we have added to the background traffic trace: EH_SENSORY for sensory data traffic, EH_IMAGE for sending medical images, EH_CONF for medical teleconferencing, as well as EH_REGISTER identifying registration

	ALL			EH_SENSORY			EH_IMAGE			EH_CONF			EH_REGISTER		
Classifier	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.
NB	0.443	0.036	0.704	0.857	0	1	1	0	1	0.929	0	1	0.966	0.051	0.532
BNET	0.93	0.009	0.952	1	0.001	0.71	1	0.001	0.8	0.929	0.002	0.65	0.994	0	1
J48	0.966	0.009	0.962	1	0	1	1	0.001	0.727	0.286	0	0.8	0.989	0	1
RT	0.964	0.006	0.965	1	0	1	0.875	0	1	0.5	0.001	0.778	1	0	0.994
SKM	0.589	0.036	0.801	0	0	0	1	0.007	0.364	0	0	0	1	0.007	0.93
Jrip	0.953	0.017	0.947	0	0.001	0	0.875	0	1	0.357	0	1	0.994	0.001	0.983
IBk	0.97	0.008	0.969	1	0	1	1	0	0.889	0.571	0	0.889	0.994	0	1
RF	0.976	0.007	0.974	1	0	1	0.875	0	1	0.286	0	1	1	0	1
MLP	0.929	0.024	0.917	0	0	0	0	0	0	0	0	0	1	0	1

information for eHealth applications. Most of the newly in-troduced applications were not recognized by the payload inspection method. For their corresponding biflows the ground truth was assigned manually. The structure of the entire trace used for experimentation is presented in Fig. 2.

For the evaluation of ML classifiers performance we have observed the True Positives (TP), False Positives (FP) and Precision metrics. TP identifies the percentage of members of the class of interest (class X) correctly classified as belonging to class X. FP specifies the percentage of members of other classes incorrectly classified as belonging to class X, whereas Precision is the percentage of instances that truly have class X among all those classified as X.

In our research on e-Health traffic classification we have used TIE (Traffic Identification Engine) [17] as the core traffic identification framework. The functionality of ML-based classifiers was incorporated with TIE through its ability to interact with a well known ML software - WEKA [18]. The information extracted from traffic traces with TIE was passed to WEKA for classification and its results were fed back to TIE, for further processing and investigation of combining techniques.

V. EXPERIMENTAL EVALUATION

The experimental setup used in our research work is pre-sented in Fig. 3. It also illustrates main eHealth applications tested in our experiments. The software was implemented on Samsung Nexus S smartphones. The monitoring station was directly connected to the access point and was intercepting traffic using packet capture software (Wireshark). Next the recorded traffic traces were analyzed to identify packet flows generated by the eHealth application.

In the experimentation phase the initial traffic trace was divided into 3 sets: training set (20%), first "test set" (40%) for initial classification, which is then a basis for the combiner setup, and the second test set (40%) on which final classifica-tion process was performed.

A. ML-based Applications Classification

For the performance evaluation of individual classifiers, the selected ML-based classifiers were identified based on the
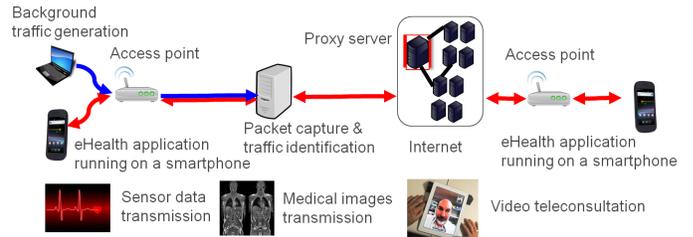

Fig. 3. Illustration of the experimental setup together with prototype applications.

state-of-the-art study. The selection was aimed to reflect di-versified classification approaches. The majority of techniques are representatives of supervised learning due to their better performance in the investigated case - Bayes Network (BNET), Naive Bayes (NB), J48 decision tree, Random Tree (RT), Random Forrest (RF), Multilayer Perceptron (MLP) based on neural networks, Jrip rule-based algorithm and K nearest Neighbors (IBk). The results present also the performance of an exemplary unsupervised classifier - Simple K Means (SKM). The initial study was performed based on the obser-vation of 24 features describing the payload sizes and inter-packet times for the first 10 packets of each flow as well as the number of packets in the flow and layer 4 protocol type (TCP or UDP). The features reflect the minimum, maximum, average and standard deviation of the flow descriptors.

Table I presents the performance results for the selected classifiers. Our particular emphasis was on the successful identification of eHealth traffic, however we also wanted to ensure the best possible accuracy for the classification of other applications. Therefore in our comparative study we present detailed results for both - the eHealth application classes and weighted average over all 13 investigated classes. The results reveal that none of the single classifiers is able to provide top accuracy levels for each of our applications of interest. Simple K Means and MLP classifiers performed very poorly for eHealth classes, with the TP close to 0 in most cases, therefore they are not considered in further evaluation. Additionally Simple K Means methods correctly classified only 39,3% of all instances and left 33,3% of instances unclassified due

Classifier	Features	No. of features
NB	1,3,4,9,11,12,19	7
BNET	2,11,12,13	4
J48	1,5,6,11,13,14, 15,21,22,24	10
RT	1,3,4,11,12,13	6
Jrip	1,6,9,11,12,13,14	7
IBk	1,4,9,11,12,13,15, 17	8
RF	1,4,6,11,12,13,14,17,20,22	10

1	l4_proto	11	up_min_plsize	19	up_min_ipt
2	up_pkts	12	dw_min_plsize	20	dw_min_ipt
3	dw_pkts	13	up_max_plsize	21	up_max_ipt
4	up_pl_pkts	14	dw_max_plsize	22	dw_max_ipt
5	dw_pl_pkts	15	up_stdev_plsize	24	dw_stdev_ipt
6	up_pl_bytes	17	up_avg_ipt		
9	up_avg_plsize	18	dw_avg_ipt		

to inability to match resulting clusters to application classes being a challenge for all unsupervised techniques. The tree-based classifiers give in general good results, however still for each of the eHealth application classes different classifier obtains the best performance. Therefore in this case methods exploiting combinations of different techniques should better meet the needs of eHealth traffic.

B. Feature Selection

The results presented in Table I were obtained on the observation of the full feature set. In practical setups, however, the observation of 24 features is not feasible and would have negative impact on the capabilities of real-time classification. Therefore we propose to limit the number of features used during classification in a way that would provide close to optimal feature set for each classifier. For this purpose the Wrapper Subset Evaluation method [19] was used together with the Best First algorithm, due to its optimization criteria, which is oriented towards maximizing the accuracy of the selected classifier. Thus different set of features is identified for each of the classifiers, as indicated in Table II. Interestingly, some of the features have not been selected for any of the classifiers - e.g. flow duration in time, downlink average payload size or standard deviation of uplink inter packet times.

For the final experiments the classifiers were trained based on the evaluation of the selected features only. This allows to limit the average number of exploited features, which varies from 4 to 10. The results obtained while exploiting only the selected features are presented in Table III. In general they are comparable to the ones obtained with full feature set, or even improved in case of Naive Bayes, however the related complexity is highly reduced. It can be observed that eHealth applications for sending sensory data and medical images are easily recognized by many classifiers with the TP up to 100% and FP close to 0. Precision values are slightly worse for the sensory data, with the minimal value of 58%. Very good results on the levels exceeding 95% for TPs are observed for

the registration services of eHealth applications. However the performance of eHealth teleconferencing traffic identification is very poor comparing to the other classes. This holds true for the results of almost all classifiers, only NB is able to reach acceptable performance. Although it could be expected that this class is incorrectly classified as SKYPE or RTP class, its mostly mistaken with BITTORRENT class. This is probably a side effect of the broad P2P definitions in the l7 filter (as indicated in previous section), which leads to the overrated P2P occurrence.

C. Multi-Classification

In the final experimentations the system was tested in a multi-classification set-up exploiting different combining techniques. There were 13 combiners [8] investigated in total, as indicated in Table IV. The results show that the best performance was obtained with the sub-types of Bayes combination rules, in particular the Maximum combiner based on the maximum rule approximation and Sum combiner based on the sum rule approximation. In this case we have again investigated the performance of eHealth traffic classification and the overall accuracy as the evaluation criteria. It is important to notice that the Average combiner (also sub-type of Bayes rule) achieved the best performance with regard to eHealth applications, however its overall accuracy was very poor. This is probably related to the structure of the investigated trace and the nature of the combiner - while the occurrences of eHealth classes are not frequent, maximizing their performance may not always lead to the growth of total performance. Moreover, in the opposite situation, most combiners aim to achieve possibly the best overall performance, which implies that they would give priority to frequent classes, since their impact on the overall performance is high. The results obtained while using the combiner allowed to increase classification performance - the TP close to 100% was achieved for e-Health sensory data, images and registration data. The accuracy for e-Health conferencing class remained relatively low, however overall it was increased to approx. 58% while the performance for other e-Health applications and for all classes remained very high, reaching 92,3% of the overall accuracy, or even up to 96,8% for the Sum combiner.

VI. CONCLUSIONS

Current networking infrastructures can support innovative eHealth applications for many users no matter whether they are at home or on the move. These applications require high QoS assurances due to the life-critical nature of the services. Unfortunately current systems cannot guarantee such QoS levels because they are unable to natively identify and tag eHealth applications. Without differentiation on the application level QoS mechanisms cannot provide appropriate scheduling and cannot reserve the required network resources.

In this paper we address this problem by proposing a multi-classification approach for eHealth applications detection and identification. In our solution we use several machine learning techniques to identify different types of traffic generated by a

TABLE III
PERFORMANCE OF ML CLASSIFIERS FOR A SELECTED FEATURE SET

Classifier	ALL			EH_SENSORY			EH_IMAGE			EH_CONF			EH_REGISTER		
	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.	TP	FP	Prec.
NB	0.879	0.024	0.881	0.857	0	1	0.875	0	1	0.857	0.001	0.75	0.966	0.011	0.844
BNET	0.937	0.014	0.938	1	0	1	1	0.001	0.667	0.214	0.003	0.273	0.989	0.001	0.983
J48	0.968	0.009	0.966	1	0.002	0.583	0.875	0	1	0.214	0	1	0.989	0	1
RT	0.964	0.008	0.963	1	0	1	0.875	0	1	0.286	0.002	0.444	0.989	0.001	0.983
Jrip	0.953	0.011	0.955	1	0.001	0.636	0.875	0	0.875	0.286	0	1	0.98	0	0.988
IBk	0.969	0.007	0.97	1	0	1	1	0	1	0.429	0	1	0.99	0	1
RF	0.972	0.007	0.971	1	0.001	0.778	0.875	0	1	0.286	0	1	0.99	0	1

TABLE IV
PERFORMANCE OF DIFFERENT COMBINING TECHNIQUES

Classifier	ALL Accuracy (TP)	SEN. TP	IMA. TP	CONF TP	REG. TP
Bayes	0.92	1	0.75	0.083	0.943
Product	0.919	1	0.75	0.083	0.943
Sum	0.967	1	1	0.5	0.989
Max	0.923	1	1	0.583	0.989
Min	0.736	1	0.75	0.083	0.94
Average	0.373	1	1	0.583	0.99
Majority Vote	0.897	1	1	0	0.98
Weight MV	0.899	1	1	0.25	0.99
Borda Count	0.9	1	1	0.083	0.98
Weight BC	0.901	1	1	0.333	0.989
Dempster-Shafer	0.953	1	1	0.167	0.989
BKS	0.938	1	0.75	0.25	0.977
Wernecke	0.934	1	0.75	0.083	0.977

mobile e-Health application. The results from each classifier are combined together to improve the overall identification accuracy. Our experimental results show, which ML classifiers and which set of features are optimal in a given case. We also investigate different combining techniques and show that the best methods for identifying eHealth applications have very high classification accuracy. Provided results and analysis prove that the proposed approach can be used as a practical solution to the problem of eHealth applications identification. It is important to highlight that the proposed solution is universal and can be also applied to identify other groups of applications not related with eHealth.

Regarding future work we plan to test our solution on other medical applications from different vendors. We will also optimize our approach for encrypted traffic and real-time operation. Future research will focus also on finding the appropriate trade-off between time, complexity and efficiency of a multi-classification system in practical conditions.

ACKNOWLEDGMENT

REFERENCES

[1] P. Koutsakis, "Guaranteed bandwidth allocation and qos support for mobile telemedicine traffic," *2008 IEEE Sarnoff Symposium*, 2008.
[2] A. Dainotti, A. Pescapè, and K. C. Claffy, "Issues and future directions in traffic classification," *IEEE Network*, vol. 26, no. 1, pp. 35–40, 2012.
[3] N. Cascarano, A. Este, F. Gringoli, F. Risso, and L. Salgarelli, "An experimental evaluation of the computational cost of a dpi traffic classifier," in *GLOBECOM 2009. IEEE*, 2009, pp. 1 –8.
[4] G. Aceto, A. Dainotti, W. de Donato, and A. Pescapé, "Portload: Taking the best of two worlds in traffic classification," in *IEEE INFOCOM Workshops, 2010*, march 2010, pp. 1 –5.
[5] T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *Communications Surveys & Tutorials, IEEE*, vol. 10, no. 4, pp. 56–76, 2008.
[6] M. Crotti, M. Dusi, F. Gringoli, and L. Salgarelli, "Traffic classification through simple statistical fingerprinting," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 1, pp. 5–16, 2007.
[7] T. Karagiannis, K. Papagiannaki, and M. Faloutsos, "Blinc: multilevel traffic classification in the dark," ser. SIGCOMM '05. New York, NY, USA: ACM, 2005, pp. 229–240.
[8] A. Dainotti, A. Pescapé, and C. Sansone, "Early classification of network traffic through multi-classification," in *Traffic Monitoring and Analysis*, ser. LNCS. Springer, 2011, vol. 6613, pp. 122–135.
[9] M. Grajzer and P. Szczechowiak, "ehealth traffic detection and classification using machine learning techniques," in *eTELEMED 2012, The Fourth International Conference on eHealth, Telemedicine, and Social Medicine*, 2012, pp. 151–154.
[10] D. Vouyioukas, I. Maglogiannis, and D. Komnakos, "Emergency m-health services through high-speed 3g systems: Simulation and performance evaluation," *Simulation*, vol. 83, no. 4, pp. 329–345, Apr. 2007.
[11] Y. Chu and A. Ganz, "A mobile teletrauma system using 3g networks." *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 4, pp. 456–462, 2004.
[12] Milwaukee School of Engineering Digital Signals and Image Database. [Online]. Available: http://myweb.msoe.edu/~martynsc/
[13] PhysioNet (PhysioBank). [Online]. Available: http://www.physionet.org/physiobank/database/#image
[14] NLANR PMA data repository (RIPE database). [Online]. Available: https://labs.ripe.net/datarepository/data-sets/nlanr-pma-data
[15] Clearfoundation, l7 filter. [Online]. Available: http://l7-filter.clearfoundation.com/start
[16] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, F. Jahanian, and M. Karir. ATLAS Internet Observatory 2009 Annual Report.
[17] A. Dainotti, W. de Donato, and A. Pescapè, "Tie: A community-oriented traffic classification platform," in *TMA*, 2009, pp. 64–74.
[18] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The weka data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
[19] R. Kohavi and G. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.