Towards the Improvement of Diagnostic Metrics
Fault Diagnosis for DSL-Based IPTV Networks using the Rényi Entropy

Angelos K. Marnerides*, Simon Malinowski†, Ricardo Morla†, Miguel R. D. Rodrigues‡ and Hyong S. Kim§
*Instituto de Telecomunicações, Faculty of Sciences, University of Porto, Portugal
E-mail:amarnerides@dcc.fc.up.pt
†INESC Porto, Faculty of Engineering, University of Porto, Portugal
E-mail:sjfm,ricardo.morla@inescporto.pt
‡Department of Electronic and Electrical Engineering, University College London, United Kingdom
E-mail:m.rodrigues@ee.ucl.ac.uk
‡Department of Electrical & Computer Engineering, Carnegie Mellon University, USA
E-mail:kim@ece.cmu.edu

***Abstract*—IPTV networks blindly rely on the adequate operation and management of the underlying infrastructure that in numerous cases is threaten by unexpected anomalous events which consequently cause QoS degradation to the end-user. Thus, it is of great importance to deploy techniques embodied with diagnostic and self-protection metrics for determining and predicting the arrival of such events in order to proactively charge defense mechanisms without the need of an exhaustive manual inspection by the network operator. In this paper we propose and demonstrate the applicability of the Rényi entropy as a useful diagnosis feature for explicitly characterizing DSL-level anomalies issued in an IPTV network of a large European ISP. It is revealed that different orders of the Rényi entropy can formulate meaningful detection and categorization of phenomena occurring on specific Digital Subscriber Line Access Multiplexers (DSLAMs) within the DSL infrastructure. Via the synergistic exploitation of the local maxima peaks generated by each Rényi-based distribution we exhibit the feasibility to extract and identify lightweight anomalies that under simple metrics cannot be detected.**

***Index Terms*—IPTV networks, DSL, Rényi Entropy, fault diagnosis**

I. Introduction

Undoubtedly, the rapid growth of the Internet's traffic volume persona is continuously challenging the operations held in Network Operation Centers (NOCs) on a world-wide scale. Naturally this new persona transformed the Internet and its services to exhibit dynamic characteristics that essentially require to be supported by mechanisms ensuring some levels of self-* properties.

In particular, services offered by ISPs such as IPTV, engage immense amounts of traffic due to user-specific requirements related with large amounts of video and audio cargo traversing the network links. On the contrary with traditional backbone networks, IPTV distribution networks exhibit unique characteristics due to their explicit QoS requirements with respect to performance, reliability and consequently maintenance. Typically small packet losses and delays caused by software/hardware failures on the backbone links may be tolerated up to a certain point but those initiated on the IPTV network would significantly degrade the video quality to the end-user. Moreover, traffic management and maintenance on the backbones is more easily predicted and easy to handle in comparison with that performed on the various infrastructure devices explicitly used by the IPTV network for broadband home users [1]. Therefore, by virtue of the sensitiveness invoked by an IPTV network it is crucial to provide metrics and strategies that can identify, predict and classify events caused by the underlay device misbehaviors. Unfortunately, current best practices employed by NOCs for infrastructure maintenance rely on a simplistic manual and empirical analysis of various events gathered by a plethora of management infrastructures (e.g. syslog) and they surely do not invoke any sophisticated metrics that can adequately adapt into the varying and unpredictable nature of such networks [1], [2].

Given the latter fact, this paper holds the incentive to improve the management domain of the underlying infrastructure by demonstrating the effectiveness and applicability of the *Rényi* entropy towards the diagnosis and root-cause identification of faulty DSLAM devices. The Rényi entropy is a useful generalization of the well known Shannon entropy and measures a signal's information in bits. To the best of our knowledge, this particular entropy formulation has not yet been thoroughly exploited within the networking literature but only in [3], [4], [5], [6] for Internet traffic classification and cryptography. Explicitly for this piece of work we measure the information invoked by the series of anomalous DSL SyncTrap events initiated on every *Digital Subscriber Line Access Multiplexer* (DSLAM) within a DSL-based network that acts as the infrastructure to an IPTV distribution network in order to identify device-related faults.

Furthermore, we illustrate that by the usage of a quite common and simple to measure "raw" statistical feature such as the inter-arrival of SyncTrap events on every *DSLAM path* (i.e. individual client), the Rényi entropy may successfully identify and diagnose a given failure. In parallel, we indicate the superiority of the Rényi entropy by comparing it with simple event-related statistics such as the mean and variance of SyncTrap inter-arrivals and the number of SyncTrap events

978-1-4673-0921-9/12/$31.00 ©2012 IEEE

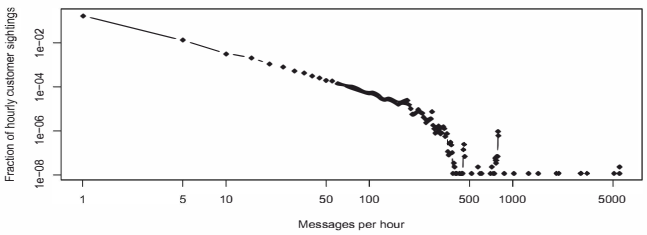

Fig. 1. Distribution of customer message counts per hour to customer sightings in the 720 hourly bins in the month.

occurring on each client. As illustrated in the dedicated results section, the different orders of the Rényi entropy may easily extract all the anomalous patterns indicated by the aforementioned statistics and in addition highlight numerous that were undetected. Therefore, due to the promising results derived out of this analysis, we argue in fair of the Rényi metric for being considered and employed in the design of future fault diagnosis schemes in IPTV networks. Via this piece of work there is also the indication that entropy-based observations only on the DSL level, may adequately express the building blocks for proactive fault awareness without the need of analyzing the application-layer traffic behavior.

Nonetheless, the remainder of this paper is structured as follows. Section II briefly presents the datasets used within our experimentation whereas section III is dedicated on illustrating the Rényi entropy as it is embodied in this piece of work. Section IV is strictly concerned with the presentation of the generated results whereas section V summarizes and concludes this paper.

II. Data Description

The data set we use in this paper has been provided by a major European commercial IP/TV service provider. We use DSL syslog logs gathered at all DSLAMs where SyncTrap messages for each time the DSL connection to any customer is lost or reestablished are present.

Our analysis depends on hourly time bins that keeps track of customer sightings alongside SyncTrap message counts in each hour. Figure 1 shows the distribution of customer message counts per hour in the 720 hours in the month (i.e. 30 days). The resulted distribution is linear in the log-log scale of the graph approximately below 200 messages per hour. Naturally, this points to a power-law relationship between hourly customer sightings and customer message count per hour, in which many customers generate a small number of messages per hour whereas a small fraction of customers generates a large number of messages per hour. This aligns with the intuition that most of the DSL customers hold a normal behavior (i.e. small number of messages per hour) and only a smaller number of customers poses anomalous statistical characteristics.

Within this work we focus on the set of high message rate customers by utilizing pairs of (customer c, hour h) observations for which customer c has more than 40 SyncTrap messages in hour h. Our dataset exhibited the majority of 92% of clients to hold more than 40 SyncTrap events in a given hourly bin. For these customers we processed a time series composed of the inter-arrival times between two consecutive SyncTrap messages related to c that occurred during hour h. Thus, the methodology illustrated next is purely dependent on the aforementioned timeseries for each customer.

III. Methodology

As introduced earlier the statistical "raw" feature that constitutes the basis in our work is the inter-arrival time of independent events for a client. There are three types of anomalous events captured by syslog for each *DSLAM path* (i.e. a single client) namely: signal degradation (i.e. type A), power cut offs (i.e. type B) and other, unknown reasons (i.e. type C). The latter class is an ambiguous category since there are certain cases where a flagged unknown event might be related with a DSLAM misconfiguration but also with a signal degradation or power cut off that was not correctly captured by the syslog management infrastructure.

Following the method explained in Section II, we have extracted from the data set for each customer c in every hour h a time-series representing the inter-arrival time between consecutive SyncTrap messages related to c and that happened during hour h. We following explain how the Rényi entropies are computed from these time-series. Let $T = t_1, \ldots, t_n$ be such a time-series. We proceed with an estimation of the Rényi entropies based on histograms. An histogram is first computed from the time-series, whose bin width complies with the Freedman-Diaconis rule [7]. According to this rule, the width of the bins for a time-series T of length n is

$$w = 2 \times IQR(T) \times n^{-1/3}, \tag{1}$$

where $IQR(T)$ is the interquartile range of T. The so-built histogram enables to model T as a realization of a discrete random variable whose possible values $x_1, \ldots, x_w$ are defined by the bins of the histogram. The probabilities $p_1, \ldots, p_w$ associated with these values are computed directly from T and the bins. According to this modeling, the Rényi entropy of order α, $\alpha \neq 1$ of T is estimated as

$$H_\alpha(T) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^{w} p_i^\alpha \right). \tag{2}$$

In this paper, we make use of the Rényi entropy of orders 2 and 3 since higher orders would lead inadequate level of sensitiveness with respect to events that have low probability such as "lightweight" anomalies that are not frequently observed. Nonetheless the Renyi entropy of orders 2 and 3, together with the Shannon entropy equal to

$$H_1(T) = -\sum_{i=1}^{w} p_i \log p_i, \tag{3}$$

can be shown to be the limit of the Rényi entropy when α tends to 1. In the following, the 1^{st} order Rényi entropy

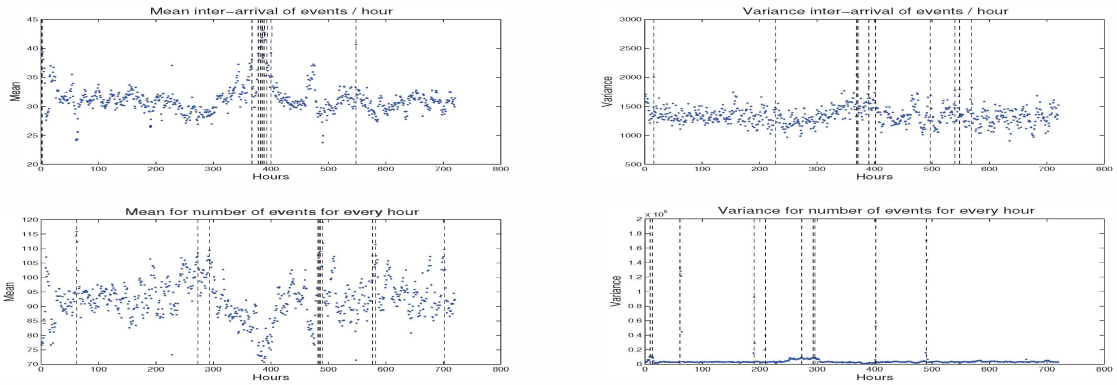

Fig. 2. Monthly mean and variance distributions for the per hour SyncTrap Inter-arrivals and number of events on every DSLAM path. Significant peaks referenced by the vertical dotted lines.

TABLE I
TOP TEN ANOMALOUS HOURS IN DECREASING ORDER AS INDICATED BY SIMPLE SYNCTRAP EVENT-BASED STATISTICS

FEATURE	*PEAKHOURS*
Avg/Inter-Arrivals	385,388,382,548,1,3,400,378,367,393
Avg/No_Events	490,62,582,702,293,486,484,273,482,577
Variance/Inter-Arrivals	497,228,369,16,371,402,390,569,548,540
Variance/No_Events	490,61,190,402,210,9,13,273,296,293

will be referred to as the Shannon entropy. Overall, we have computed for every hour and every client the Rényi entropies of order $\alpha = 1$, $\alpha = 2$ and $\alpha = 3$. Apart from computing the Rényi entropies and for data dimensionality redundancy issues we aggregated the Rényi values obtained for every hour and exploited them as random variables by observing the mean and variance distributions as well as their corresponding 1^{st} order differenced series. Consequently, these observations settled as the basis for constructing conclusions regarding faulty DSLAM paths as well as for main DSL multiplexers.

IV. RESULTS

The formulations introduced in Section III were employed on the datasets described in Section II. Under the intention of coherently presenting our resulted outcomes, this section firstly illustrates the observations derived by analyzing basic statistics extracted by simple event-related features and then exhibits the benefits attained via our proposed Rényi-based methodology. The rationale behind the structure of this section relates with the comparison of these features and the justification of our argument with respect to the autognostic properties that our scheme can facilitate.

A. Diagnosis under common features

Intuitively and as commonly practiced by network operators, anomalous hourly patterns can be reasonably identified by simply looking at the volume of SyncTrap events initiated on the numerous DSLAM paths (i.e. clients). At the same time, a more fine grained analysis can also be allowed by exploiting some statistical properties of these events such as their inter-arrival patterns that we also use in this work. Figure 2 shows the monthly distribution of the average and variance values for the SyncTrap inter-arrival times (top plots) and the number of SyncTrap events measured in all active DSLAM paths (bottom plots). Indeed, through the plots provided it is clearly evident that there are several anomalous hourly bins manifested by delayed or accelerated SyncTrap events or by heavy event hitters on particular clients.

As anticipated, the average and variance distributions would have conflicting outputs with respect to their local minima and maxima points and naturally they would indicate different anomalous hour bins. This expected observation is firmly justified by Table I where the top ten hourly peaks detected by each event-related statistical feature are presented. By considering the peaks of Table I it is clearly noticed that all four metrics cannot be fully agreeable thus they should be collaboratively used. Nevertheless, these four metrics cannot adequately enlighten a network operator with respect to the explicit type of the SyncTrap events that caused these anomalous patterns. In addition and as we show in the following subsection, these metrics cannot guarantee that they can fully detect and categorize the anomalies present within our datasets.

According to our cross-validation accomplished by mining back to the real datasets, the majority of peaks related with the mean volume and variance of SyncTrap messages, were mostly due to immense amounts of type C events where in many cases they were initiated on the same time instant (i.e. inter-arrival time was equal to 0). On a much smaller number of cases and particularly for around 1% of all the observed local maxima the anomalous DSLAM paths had a large amount (e.g. $500 <$ events < 1800) of type A events (i.e. signal degradation). Unfortunately none of the analyzed peaks corresponded to any type B events which indicate power cut offs.

Strengths and pitfalls were also exposed by our further analysis on the average/variance inter-arrivals distribution. Similarly with the volume-based investigation discussed earlier, it was realized that around 90% of the analyzed hour

bins were directly related with unknown reasons (i.e. type C events) whereas 3% were due to power cut-offs and 7% initiated by signal degradations. However, on the contrary with the volume-based analysis, the inter-arrival feature could characterize better the type C events. There were many DSLAM paths contributing to the overall average/variance of the hour bin that in reality had an extremely small number of type-C events (e.g. 40 < events < 100) but with extremely large inter-arrival times between these events (e.g. $20sec$ < int_time < $2min$).Nonetheless, the most intriguing fact from the inter-arrival investigation was related with the shape of the variance distribution (Figure 2 bottom right plot) which in practise has shown 5 significant peaks where the rest of considered peaks laid around the mean variance value. Interestingly enough these 5 peaks were issued by single "bursty" DSLAM paths that possessed more than 5000 type C messages. Likewise, the peaks obtained above the mean variance value were also due to type C events present on single DSLAM paths holding a much smaller number of events. Overall, the variance analysis has enlighten the relationship of the inter-arrival variance and the number of events, but still it could not adequately map the different types of anomalous SyncTrap events.

Summary: The experimentation presented above has identified the strengths and weaknesses posed by simple features that can easily be used by a network operator. It has been demonstrated that a volume analysis based on the aggregate measurement of per DSLAM path events would only reveal bursty clients labeled with type C (i.e. unknown) events. On the other hand, an analysis based on the inter-arrival patterns of SyncTrap events would provide more fine-grained but not ideal diagnostics with respect to the exact root-cause of an anomalous event. Both features could not detect all the anomalous patterns residing in the dataset. The outcomes of this small and simple analysis emphasize the importance for considering new metrics that can adequately detect and diagnose SyncTrap events.

B. Rényi-based Analysis

The features described earlier seemed unable to fully perform a granular fault diagnosis, thus we continue by illustrating our Rényi-based solution that overcomes their limitations. As already mentioned we follow a simple analysis and we treat the resulted per/client Rényi entropies of orders 1,2 and 3 as random variables and observe their monthly mean and variance distributions in order to identify anomalous peaks.

Figure 3 shows the monthly distribution of the mean and variance entropy (i.e. 720 hours) of the captured trace for all the three Rényi orders. From a general point of view it can be fairly said that the plots corresponding to the overall entropies mean (left column plots) and variance (right column plots) hold a similar shape finding themselves agreeable with respect to the most dominant peaks. A detailed analysis achieved after back-tracking with the real datasets revealed that collaboratively all Rényi orders from their mean and variance perspective identified all the anomalous hourly bins detected by the features described in Section IV-A and they have further spotted new significant peaks.

TABLE II
TOP TEN ANOMALOUS HOURS IN DECREASING ORDER AS INDICATED BY THE MEAN AND VARIANCE DISTRIBUTIONS OF THE RÉNYI ENTROPIES

$FEATURE$	$PEAK HOURS$
$H_1(T)$ *Mean*	1,388,3,383,17,380,548,365,367,15
$H_1(T)$ *Variance*	657,608,241,259,653,714,417,630,86,659
$H_2(T)$ *Mean*	18,16,380,384,1,386,388,365,3,23
$H_2(T)$ *Variance*	241,608,653,418,229,261,714,423,259,568
$H_3(T)$ *Mean*	16,18,381,378,385,23,1,365,412,473
$H_3(T)$ *Variance*	418,241,229,261,608,376,244,653,356,636

As shown by Table II a substantial percentage of the most significant peaky hours presented earlier (Section IV-A, Table I) have been successfully noticed by the entropy-based statistics. The reason of some peaks from Table I not being present in Table II relates with the fact that the entropy-based statistics considered them as insignificant peaks and ranked them into lower impact ranks. Indeed, abnormal hourly bins (e.g. 490^{th},497^{th},385^{th} bins) pointed as critical by the simple features described in Section IV-A were categorized to a lower rank whereas other earlier unnoticed bins such as those for the 18^{th}, 16^{th},381^{st} hours were assigned as highly critical from an entropy point of view. Leaving aside this ranking sensitiveness offered by the Rényi-based features, it was in addition appealing enough to observe that local maxima and average points within the mean and variance distributions on all three orders could map peaks to specific types of SyncTrap events on particular DSLAM paths. Moreover, it was generally concluded that all the three Rényi orders assigned volume-based events close to either their mean or variance average range of values. Under this capability, they accommodated a much detailed insight with respect to "lightweight" anomalies that are not easily observed from a typical volume-based analysis.

In more detail, we have observed that 80% of the hourly bursts summarized by the Shannon entropy mean were caused by no more than 3 customers with a small number (i.e. 40 < events < 80) of type C events (i.e. unknown reasons) that exhibited a medium-sized range of inter-arrivals where 15sec < int_time < 35sec. Similarly, the variance-related peaks allocated for the Shannon entropy were up to 85% related with type C events initiated by more than 5 clients with less than 100 SyncTrap events triggerred within large inter-arrivals (e.g. 25sec < int_time < 2min).

Nonetheless, the most attractive results regarding the categorization and diagnosis of SyncTrap events were demonstrated by the findings of the 2^{nd} and 3^{rd} order Rényi entropies. In particular, the mean distribution of the 2^{nd} order Rényi entropy could by far associate the detected hourly peaks with signal degradation occurrences (i.e. type B events). On a 90% accuracy percentage the peaks indicated were caused on less than 3 clients per hour who experienced a small number of signal degradation events (40 < events < 80) on their corresponding DSLAM path within really large inter-arrivals. Apart from identifying the aforementioned type of clients, the

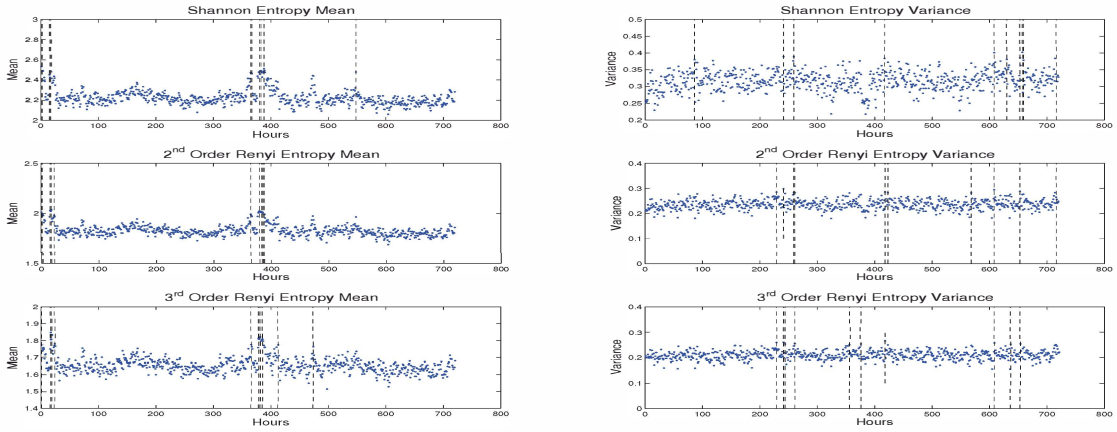

Fig. 3. Monthly distribution for the mean and variance of different entropy orders for per/client SyncTrap inter-arrivals. Significant peaks referenced by the vertical dotted lines.

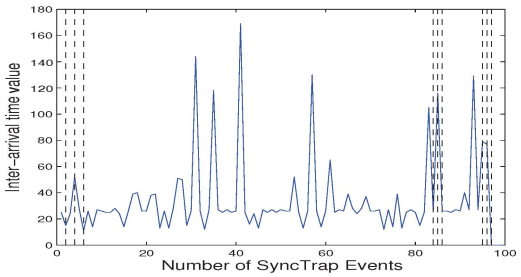

Fig. 4. An example of the inter-arrival timeseries for a client in the 384^{th} hour bin that experienced simultaneously a signal degradation and power cut-offs(black dotted vertical lines).

outcomes of the 2^{nd} order Rényi also indicated some anomalies that we didn't have a priori knowledge about them. Within several of the top ranked hourly peaks such as the 380^{th} and the 384^{th} bins the high entropy mean was caused due to DSLAM path that contained two types of SyncTrap events. There were more than 5 clients who experienced both type A and type B events (i.e. signal degradation and power cut-off) in the aforementioned bins. Their generic characteristics are strongly similar with the exemplar plot of Figure 4 where the interarrival timeseries of an anomalous client of this type is given. Via this plot it is quite evident that the inter-arrival values of power-cut offs which are mapped as the black-dotted points exhibit a varying nature in their values. Based upon this fact, we consider as logical the inbaility of the analysis described in Section IV-A to detect any of these phenomena.

However, we feel the essence to mention that the 2^{nd} order Rényi mean distribution itself did not fully capture all of the anomalous clients but it was complemented by the resulting outcomes of the 2^{nd} order Rényi variance distribution. In addition, the variance distribution also pointed to other lightweight signal degradation phenomena in several DSLAM paths that volume-wise seemed insignificant as well as low-volume clients that experienced DSL disconnections due to unknown reasons (i.e. type C events). The greatest number of the latter class of events which attained an 80% of the detected peaks demonstrated extremely small inter-arrival values of less than 15 seconds whereas the former had large inter-arrivals of more than 50 seconds.

On the contrary, the 3^{rd} order Rényi findings could not relate with any of the type A or type B events but they produced a good accuracy level on the diagnosis of type C events issued under an even smaller range of inter-arrival values of less than 10 seconds. Under both the mean and the variance viewpoint, the local maxima points of each generated distribution could synergistically expose type C anomalous patterns that could have not been discovered by any of the lower order entropies. Furthermore, hourly bins around the average on both distributions enabled a clearer categorization of DSLAMs that only experienced type B events (i.e. power cut offs).

Summary: This subsection has demonstrated the benefits offered by a collaborative analysis of the first three Rényi entropy orders for fault diagnosis purposes. In general, the experimentation conducted has empowered the argument with respect to the essence of these entropy-related metrics for identifying lightweight anomalies that cannot be diagnosed under simple features as commonly used by network operators. By contrast with the findings in Section IV-A it was shown that the Rényi-based mean and variance distributions could easily locate the peaks indicated by simple features and further extract many more anomalous patterns. In particular it was illustrated that the Shannon entropy which in practice denotes the 1^{st} order Rényi entropy enabled the mapping of type C events in the majority of its peaks whereas the synergistic analysis under the 2^{nd} and the 3^{rd} order Rényi could adequately locate many instances of DSLAM paths that experienced type A and B events.

TABLE III
TOP TEN ANOMALOUS HOURS IN DECREASING ORDER AS INDICATED BY THE DIFFERENCED MEAN AND VARIANCE DISTRIBUTIONS OF THE RÉNYI ENTROPIES

FEATURE	*PEAKHOURS*
Diff $H_1(T)$ *Mean*	376,14,399,547,441,320,186,604,92,69
Diff $H_1(T)$ *Variance*	87,240,658,618,77,57,656,468,604,212
Diff $H_2(T)$ *Mean*	376,14,441,186,320,140,497,92,69,130
Diff $H_2(T)$ *Variance*	635,57,87,240,658,33,618,308,567,77
Diff $H_3(T)$ *Mean*	497,376,441,14,186,140,130,92,320,69
Diff $H_3(T)$ *Variance*	635,57,87,308,240,417,260,33,90,186

C. Complementary Fault Diagnosis

Considering the unpredictable persona that SyncTrap events may expose, our investigation went a step ahead in order to identify the most peaky hour transitions that may be initiated either from an event volume perspective or by the intrinsic inter-arrival characteristics of a particular type of event on a range of DSLAMs. The aim was to locate incremental patterns throughout the time domain in order to filter out normal activity and further pinpoint extremely anomalous hourly bins. By virtue of the insightful results obtained by the entropy metrics discussed earlier, we simply employed 1^{st} order differentiation to all of the Rényi mean and variance distributions and examined the value-wise difference between consecutive hourly bins.

Table III demonstrates in a decreasing order the most high impact hour bins that exhibited an extremely large increase of either their entropy mean or variance distribution. As anticipated, some selected transition peaks were spotted by all the differenced Rényi orders but in many other instances the results were not agreeable. Moreover, a great portion of the selected peaks was not also visible through the experimentations performed in sections IV-A and IV-B. Truly, all of the indicated transitions were related with abnormalities including both volume-wise or inter-arrival incremental/decremental behaviours on several DSLAM paths. The majority of SyncTrap volume-wise transitions were spotted by explicitly examining the differenced orders of the entropy variance distributions of all order whereas inter-arrivals mostly dealing with type A and type B events were easily spotted by the differenced series of the 2^{nd} and 3^{rd} order Rényi mean distributions. However, tremendously big peaks were never issued by power failure events (i.e. type B events) but rather up to a 95% were due to signal degradations (i.e. type A events).

Summary: The analysis discussed in this subsection has empowered upon the benefits embodied within a Rényi-based fault diagnosis methodology. It was evidenced by Table III that when the differenced series of the entropy-based measurements are examined it is feasible to identify anomalous hourly transitions that cannot be reasonably detected by static simple features as presented by Table I in SectionIV-A. Naturally, such outcomes may be exploited for fault diagnosis forecasting purposes since the statistical predictability disclosed within the Rényi entropy formulation can ensure up to a great extent an accurate diagnosis of abnormalities.

V. CONCLUSION

The emerging Internet service technologies offered today by ISPs involve highly challenging tasks for network operators regarding the identification of faults occurring either on a local or a network-wide level. It is of crucial and critical importance for any operator to be equipped with mechanisms embodied with powerful metrics that can adequately facilitate fault diagnosis solutions. In this piece of work we have emphasized and introduced the beneficial abilities of the Rényi entropy within the explicit scenario of fault diagnosis on DSL-level anomalies for IPTV networks.

Based upon a single network-based feature such as the inter-arrival of SyncTrap events on DSLAM paths it was demonstrated that the various orders of the Rényi entropy may collaboratively constitute robust conclusions for identifying a failure over a DSL line for a designated client. For this particular scenario, it was shown that these entropy-based metrics are by far more superior than simple statistics which are commonly used by network operators. By contrast with simple descriptors such as measurements of the volume and the inter-arrival of SynTrap events, it was presented that the Rényi-based estimations may easily extract several "lightweight" anomalies holding insignificant volume-wise persona. In addition, this work has also illustrated that via a simple analysis of the differenced entropy orders it is possible to pinpoint large hourly anomalous transitions irrespective of the volume-wise characteristics of these bins. In general, we argue that the promising results presented in this work can definitely broaden the horizons within the realms of development of fault diagnosis tools for any networked scenario.

ACKNOWLEDGMENT

This work was funded by the FCT (Portuguese Science & Technology Agency) project NeTS:CMU-PT/RNQ/0029/2009 under the CMU-Portugal research collaborative scheme. The authors would like to thank FCT for its vital support.

REFERENCES

[1] A. A. Mahimkar, Z. Ge, A. Shaikh, J. Wang, J. Yates, Y. Zhang, and Q. Zhao, *Towards automated performance diagnosis in a large IPTV network*, in Proceedings of ACM SIGCOMM 2009, pp. 231-242.

[2] A. A. Mahimkar, H. H. Song, Z. Ge, A. Shaikh, J. Wang, J. Yates, Y. Zhang and J. Emmons, *Detecting the performance impact of upgrades in large operational networks*, in Proceedings of ACM SIGCOMM 2010, pp. 303-314

[3] A. K. Marnerides, *On Characterization & Decomposition of Internet Traffic Dynamics, Ph.D Thesis*, Department of Computer Science and Communications, Lancaster University, UK, September 2011

[4] A. K. Marnerides, C. James, A. Schaeffer,Y. Sait, A. Mauthe and H. Murthy, *Multi-level Network Resilience: Traffic Analysis, Anomaly Detection & Simulation*,in ICTACT Journal, Special Issue on Next Generation Wireless Networks and Applications, Vol 2, Issue 2, 2011

[5] Z. Chen and C. Ji, *An Information-Theoretical View of Network-Aware Attacks*, IEEE Trans. Information Forensics and Security, Vol. 4, No. 3, pp. 530-541, Sept. 2009

[6] C. Cachin, *Entropy measures and unconditional security in cryptography*, Ph.D thesis, Swiss Federal Institute of Technology, Zurich, 1997.

[7] D. Freedman and P. Diaconis, *On the histogram as a density estimator*, L2 theory. Probability Theory and Related Fields (Heidelberg: Springer Berlin) 57 (4): 453Ű476. ISSN 0178-8051, 1981