

A Passive Measurement Tool for Monitoring Wireless Packet Network Performances

Luigino Benetazzo¹, Claudio Narduzzi¹, Paolo Attilio Pegoraro¹, Ronny Tittoto²

Department of Information Engineering, University of Padova,
via Gradenigo 6/a - 35131 Padova, Italy
Phone: +39 049 8277760, Fax: +39 049 8277699

E-mail: ¹{luigino.benetazzo, claudio.narduzzi, paoloattilio.pegoraro}@unipd.it, ²ronny.tittoto@tiscali.it.

Abstract – Continuous monitoring of performances becomes necessary to detect critical situations in the radio interface of digital wireless telephone networks. This paper presents a measurement tool that analyzes TCP-level packet traffic and detects the effects of variations in the radio channel conditions of a mobile network. The analysis algorithm considers the energies of wavelet coefficients of the aggregated traffic process at a suitably selected number of time scales. Wavelet energy variations are detected by a statistical test from which a score is calculated. This indicator reveals changes in the performances of the monitored network.

Keywords – passive measurement, wireless network monitoring, wavelet-based traffic analysis.

I. INTRODUCTION

Digital wireless telephone networks will increasingly rely on measurements to operate efficiently. Packet-based data services are growing in importance, requiring the provision of adequate quality for different applications sharing the same network. An over-provisioning approach is hardly possible in this context, because radio resources are scarce and precious. Furthermore, predefined and static rules are unsuitable for planning future networks, since data volumes will be much variable; so-called capacity hotspots might be generated, whose location and load would vary as the population of users changes or moves. In the light of these developments, a measurement-oriented point of view would provide the required flexibility and the ability to adapt to quality of service requirements.

While performances of traditional networks could be summarized by a few simple indicators (roughly speaking, bit error rate is the key parameter), nowadays packet-switching data networks require more complex performance metrics for a wider variety of purposes: verification of service level agreements (SLAs), accounting and billing, resource management and traffic engineering. Consequently, providers need new and more sophisticated analysis tools to gather real-time information about network performances.

Passive measurements, based on the non-intrusive observation of normal user traffic, seem to be most appropriate in the framework of wireless mobile networks. In fact, these networks are subject to strict resource constraints and the increase of traffic due to the injection of markers or artificial traffic would often not be tolerable. Passive observation of traffic

requires suitable measuring equipment: at the physical level this would be represented by measuring receivers, but traffic is better analyzed at higher level. Then, measurements no longer deal with physical quantities representing signals, but rather involve the observation of data packets flowing through the network. Instruments commonly employed for this purpose may be as complex as a communications protocol analyzer, but much simpler devices, like a packet probe, can be built around less expensive network interface card.

An important issue is to decide at which level of the protocol hierarchy the analysis should take place. Measurement at the transport level shows to have advantages, as it allows to circumvent the need for coordinating several dedicated approaches, which would be typical for a network-level monitoring framework. In fact, analysis at the transport level is independent of the different technologies underlying packet delivery and is based on the observation of packet traffic flows and of their aggregations in certain network nodes. On the other hand, complex analysis of the observed quantities may be needed to infer meaningful performance indications.

Since wireless networks are particularly affected by radio condition variations, continuous monitoring becomes necessary to detect critical situations in the radio interface. This paper presents a measurement tool that analyzes TCP-level packet traffic and detects the effects of variations in the radio channel conditions of a mobile network.

To perform this task efficiently, it is preferable to observe traffic flowing through the wired link that runs between a radio base station and the fixed, broadband data network. The monitoring instrument acquires traces containing just basic information about packets flowing through the link, which are condensed into a record of data representing the random process of aggregated traffic flowing through the observation point. The proposed tool is based on the use of the wavelet transform [1] for the analysis of aggregated traffic.

The analysis algorithm considers the energies of wavelet coefficients at a suitably selected number of time scales. In the following sections it will be shown that their values are related to network performances, namely, to a radio link parameter called the block error rate, which significantly affects traffic capacity. To provide a ready indication of potential malfunction-

tions or configuration problems, calculated energy values have been combined into a score, i.e., an index evidencing trends towards improvement or degradation in network performances. Statistical properties of the proposed score are considered and carefully analyzed.

Finally, application to the monitoring of TCP performances in a realistic third generation wireless network topology [2] is discussed in greater detail and results obtained with the proposed statistical wavelet-based indicators are presented, to show that the tool is effective as a means to discriminate radio conditions and track their effects.

II. MEASUREMENT ON AGGREGATED TRAFFIC BY WAVELET ANALYSIS

In the fixed segment of a mobile network, links can be monitored passively by means of special probes that gather information on packets and protocol headers. The resulting traces contain information about both the timing and the size of packets flowing through the observed point. The traces are then processed to obtain the relevant measurement information. Fig. 1 presents a simple scheme of a typical topology in a mobile network scenario, highlighting also the measuring point.

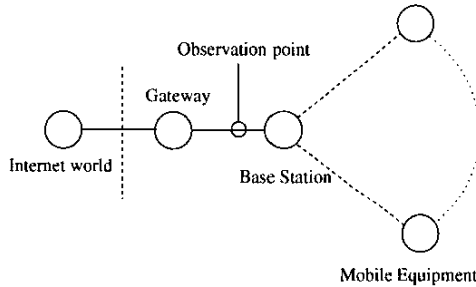


Fig. 1. Topology of the analyzed mobile network

The analysis considered in the paper deals with aggregated traffic, i.e., the total traffic observed during a succession of time intervals. Aggregated traffic, with aggregation time step T , is formally defined as a time series $X_T(n)$, each element corresponding to the number of packets (or bytes, depending on the aggregation criterion) observed in an interval $[nT, (n+1)T)$. The monitoring procedure is based on the measurement of aggregated traffic directed toward a base station. A sliding window of duration D determines the measurement time; the number of samples is therefore $N = D/T$.

A. Wavelet analysis

The wavelet transform is a very effective tool in the analysis of signals in different fields of research such as geophysics, signal processing, pattern recognition [3]. The discrete wavelet decomposition (DWT) represents a given signal or process $f(t)$ in terms of the shifted and dilated versions of a prototype

bandpass “wavelet” function $\psi_0(t)$ called “mother wavelet” and shifted versions of a low-pass “scaling” function $\phi_0(t)$ [4]. With particular choices of wavelet and scaling functions the atoms:

$$\begin{aligned}\psi_{j,k}(t) &= 2^{-\frac{j}{2}} \psi_0(2^{-j}t - k) \\ \phi_{j,k}(t) &= 2^{-\frac{j}{2}} \phi_0(2^{-j}t - k)\end{aligned}\quad (1)$$

form an orthonormal wavelet basis and $f(t)$ can be expressed in the following way:

$$f(t) = \sum_k c_{J_0,k} \phi_{J_0,k}(t) + \sum_{-\infty}^{J_0} \sum_k d_{j,k} \psi_{j,k}(t), \quad (2)$$

where the first term represents an approximation of $f(t)$ at scale 2^{J_0} and the remaining part is the contribution of irregularities at every scale 2^j . The result of the DWT is the set of coefficients $c_{J_0,k}$ and $d_{j,k}$, obtained by projection of $f(t)$ on the chosen basis.

The DWT can be efficiently implemented by means of a recursion of lowpass and highpass filters alternating with decimation by a factor 2 [3]. This takes the form of a digital filter tree (see [4] for details), which represents an efficient $O(N)$ algorithm.

The wavelet analysis of aggregated traffic is based on the computation of the DWT of the traffic process $X_T(n)$. This yields detail coefficients $d_X(j, k)$, where j is the scale index, corresponding to a time scale $2^j T$, and k is the translation index. The detail coefficient $d_X(j, k)$ measures the process content around time $k2^j T$ and in the neighborhood of frequency $2^{-j} f_0$, where f_0 is the “center” frequency of the mother wavelet ψ_0 .

B. Second order analysis

Detail coefficients highlight variability in the behavior of $X_T(n)$ at different time resolutions and help identify significant features in the traffic pattern. To investigate second order properties of $X_T(n)$, the energy functions E_j can be introduced:

$$E_j = \frac{1}{N_j} \sum_{k=k_0}^{k_0+N_j-1} |d_X(j, k)|^2, \quad j = 1, 2, \dots, J, \quad (3)$$

where N_j is the number of coefficients at scale j and J is the coarsest scale index. E_j can be interpreted as the mean energy associated to irregular components in the behavior of $X_T(n)$ at scale $2^j T$. A small energy at this scale points to the presence of regularities or periodical patterns in the time series. A representation of energy versus scale in a log-log plot (LogScale diagram) is commonly used for this kind of analysis [5], [6].

In the context of performance evaluation [1] and troubleshooting detection it is important to determine which time scales are relevant for the observation of the traffic process. If attention is concentrated on the description of TCP dynamics, analysis must be focused on time scales typical of its feedback mechanisms, such as self-regulation and synchronization. In

fact, degradation of network performances is often reflected into greater variability of TCP timing parameters. A suitable choice is to consider the scales related to typical feedback mechanisms, such as Automatic Repeat reQuest (ARQ). The time scales of the TCP Round Trip Time (RTT) are of particular interest: within that range the aggregate traffic under good radio channel conditions presents regularities which are evidenced by a minimum of the energy function E_j [7]. This tends to disappear when performances are degraded.

III. NETWORK MONITORING ALGORITHM

The monitoring procedure consists of comparing the energies E_j obtained in different time windows of equal duration D . Starting from $N = 2^J$ samples, the digital filter tree DWT algorithm yields J scale levels with decreasing resolution (due to the decimation process the number of DWT coefficients also decreases from 2^J to 1 per scale). For each monitored window the associated scale energies E_j for $j = 1, \dots, J$ are computed and stored so that network behavior can be tracked. To detect radio channel variations and raise an alarm in case of badly degrading conditions, energies at each scale need to be compared with their correspondents in another temporal window. Changes in scale energies must be detected and, for this purpose, a specific algorithm is required. In fact, a traditional hypothesis test cannot be applied because of the non-normal distributions of the wavelet coefficients at each scale, which make computation of confidence intervals for the energy values very difficult. A better way to compare the energies is to apply non parametric hypothesis tests and in this work the use of Miller [8] and Moses tests [9] is proposed. Their features are briefly described in the following subsection.

A. Non parametric hypothesis tests

To compare statistical samples (and in particular their means or variances) whose distribution is unknown non parametric tests are necessary. These tests are based on minimal assumptions, do not require a priori models for the data under study and are therefore called "distribution-free" tests. The aim of this study is to assess changes in wavelet coefficients energies at each scale of analysis.

Wavelet coefficients $d_X(j, k)$ for a given scale j are realizations of a zero mean stochastic variable. It can be seen from (3) that the corresponding energies can be interpreted as a variance estimator for this variable. Then, instead of comparing energies, the sets of wavelet coefficients determined in two different observation windows are considered as statistical samples with the aim of comparing their variances. The Moses and Miller tests were therefore chosen, among several possibilities, because they are designed for testing dispersion. Both tests do not require the means of the two samples to be equal.

The null hypothesis is the equality of dispersion parameters between two statistical samples: $H_0 : \sigma_1 = \sigma_2$, or equivalently: $H_0 : \gamma^2 = \left(\frac{\sigma_1}{\sigma_2}\right)$. The alternative hypothesis has

been assumed to be the increase of dispersion: $H_1 : \sigma_1 > \sigma_2$ (one-sided upper-tail test). Every statistical test gives as result a probability index, called *p-value*, that allows a decision to be made. When the *p-value* tends to zero the null hypothesis is unlikely.

The Moses test is performed by the following three steps [9]:

1. The original statistical samples to compare, A and B , are both randomly divided into m subsets of equal size k .
2. For each subsample the sum of squared deviations from the empirical mean is computed. In this way two new statistical samples, C and D , are formed by the m values obtained from A and B , respectively.
3. The Mann-Whitney-Wilcoxon non parametric location test is applied to the two dispersion samples C and D .

The Mann-Whitney-Wilcoxon test is performed by merging C and D into single ordered set and assigning a positional index to each value in two sets. The positional indexes of one of two sets, e. g., C , are summed up to get a "rank sum score" S . Under the null hypothesis the statistic S is, in the large-sample approximation, a normal variable with mean and variance depending only on the sample size m . Its normalized version z allows then to compute the *p-value* as $1 - \Phi(z)$ where Φ is the cumulative normal distribution.

The Miller test is based on the so-called jackknife procedure [8] which involves a more complex sequence of steps. From the original sets A and B (of size n and m respectively) the statistic Q is computed as follows:

1. Compute the averages and variances for set A :

$$\bar{A}_i = \sum_{s=1}^{i=n} \frac{A_s}{n-1} \text{ and } D_i^2 = \sum_{s=1}^{i=n} \frac{(A_s - \bar{A}_i)^2}{n-2},$$

likewise, the set B yields: $\bar{B}_i$ and E_i^2 .

2. Define $S_i = \ln D_i^2$ and $T_i = \ln E_i^2$. In addition let:

$$S_0 = \ln \left[\sum_{s=1}^n \frac{(A_s - \bar{A}_0)^2}{n-1} \right] \text{ where } \bar{A}_0 = \sum_{s=1}^n \frac{A_s}{n},$$

T_0 is obtained in the same way for sample B .

3. Apply the "jackknife" procedure by computing the following variables:

$$M_i = nS_0 - (n-1)S_i, \text{ for } i = 1, \dots, n$$

$$N_j = mT_0 - (m-1)T_j, \text{ for } j = 1, \dots, m$$

$$\bar{M} = \sum_{i=1}^n \frac{M_i}{n} \text{ and } \bar{N} = \sum_{j=1}^m \frac{N_j}{m},$$

$$V_1 = \sum_{i=1}^n \frac{(M_i - \bar{M})^2}{n(n-1)} \text{ and } V_2 = \sum_{j=1}^m \frac{(N_j - \bar{N})^2}{m(m-1)},$$

4. Finally, the statistic Q can be computed:

$$Q = \frac{\bar{M} - \bar{N}}{\sqrt{V_1 + V_2}},$$

Under the large-sample approximation, the statistic $Q \sim \mathcal{N}(0, 1)$ if the null hypothesis H_0 is true. Then, like in the Moses test, the p-value can be easily obtained as $1 - \Phi(Q)$.

Thanks to the jackknife procedure, the Miller test allows to better deal with repeated values within each sample, an important issue for the analysis of the packet traffic process.

B. Considerations on test results

In the context under study an increase in wavelet coefficients energy at a given scale j can be seen as a loss of regularity in the traffic behavior induced by TCP self-regulation. This is a footprint of worsened radio conditions and is revealed by the one-sided test proposed above when a p-value p_j close to zero is obtained.

p-values can also be used as a quantitative index of how far the data are from satisfying the null hypothesis. Since a number of different scales are considered, it is possible to compute a parameter that summarizes the test results. A weighted summation of transformed p-values p_j is proposed as a final score:

$$S = \sum_{j=j_1}^{j=j_2} g(p_j)w_j, \quad (4)$$

where w_j are weighting coefficients and $g(\cdot)$ is a suitable monotonic function (in the following $\log_{10}$ is used). j_1 and j_2 are the limits of the observed scale range. Since a sliding analysis window is employed, the behavior of this synthetic indicator can be followed over time and helps to track radio channel variations.

IV. RESULTS

The proposed measurement tool has been developed and tested, at first on artificially generated traffic obtained via ns-2 [10] simulations. To obtain realistic traffic simulations, UMTS Radio Link Layer (RLC) and Medium Access Control (MAC) protocols were implemented, carefully following 3GPP standards [11], [12] and ensuring that all aspects that have a potential impact on performances were correctly implemented. Particular emphasis should be placed on the ARQ mechanism of the RLC layer. In fact, an IP packet is fragmented in a number of blocks that are passed to the MAC and physical layer for transmission. Each block received by the listening RLC produces an acknowledgment block that is sent back to the sending RLC. Transmission blocks can be affected by errors in the radio channel, in which case they are retransmitted until

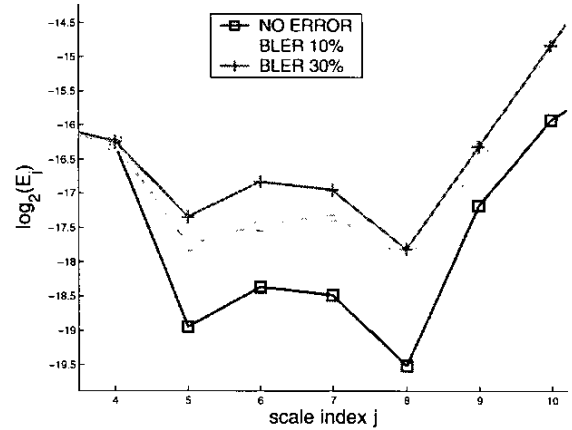


Fig. 2. LogScale diagram for three different radio conditions

the corresponding ACK is received¹. So a TCP packet may require more time to be completely transmitted in case of radio losses and RTT values become larger and more variable.

An extensive validation of the UMTS simulation modules has been carried out, after which a realistic network topology (see Fig. 1), comprising ten mobile terminals in a web-browsing framework, was set up.

Several simulations have been performed, with changing radio conditions. The channel error model was a two-state model where one state is error free and the other is characterized by a BLock Error Rate parameter, that will be referred to as BLER in the following. The model also includes a matrix of transition probabilities that were obtained from literature findings [13].

The simulator was first employed to produce throughput-versus-BLER diagrams. Their analysis provided an indication of representative BLER values, that were used to set the simulations so that different network efficiency conditions could be analyzed.

Fig. 2 shows a LogScale diagram associated with three different BLER conditions, obtained with an aggregation rate $T = 10\text{ms}$ and a window of $N = 2^{16}$ samples. Reported results are relative to BLERs equal to 10% and 30% and are compared with a no-error condition. Scale indexes from $j = 4$ to $j = 10$ were chosen for the analysis as they are the most relevant for TCP self-regulation phenomena and can be identified via a very simple RTT estimator. They correspond to times ranging from 160ms to 5s. The diagram gives a qualitative idea of how an energy pattern can change as a consequence of radio channel variations. It can be noticed that virtually no variation occurs for $j = 4$, and this should be reflected in the outcomes of the non-parametric tests.

Moses and Miller non parametric tests have been applied to simulation data where a given radio channel condition was set and kept constant throughout the time window. BLER values

¹ The maximum number of possible transmissions of the same block can be configured by the RLC layer parameter MaxDAT

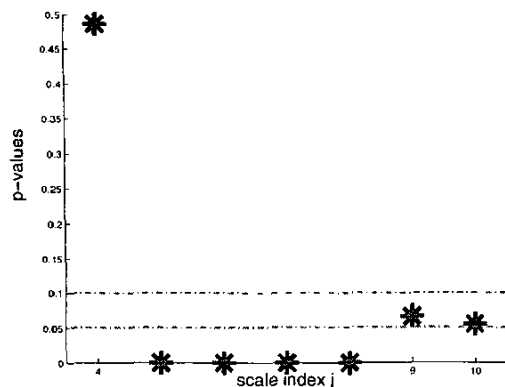


Fig. 3. Miller test p-values in the comparison of 10% BLER data versus 0% BLER ones (vertical log-scale).

ranging from 0% to 30%, like in Fig. 2, were considered. Fig. 3 shows the p-values resulting from the Miller test when the case BLER = 10% is compared to the no-error case. In statistical tests, a threshold p-value is considered in order to make a decision (usually 0.05 or 0.1 as shown with dash-dotted lines in Fig. 3). Then, the null hypothesis would be surely rejected at scales from 5 to 8, while scales 9 and 10 are on the edge of rejection. It is possible to see from Fig. 2 that the only scale that doesn't reveal a changing condition is $j = 4$ and this is reflected by the high p-values in Fig. 3 at the same scale.

The score introduced in the previous section summarizes in a quantitative way the variations and provides a simple and meaningful indicator. Scores S in Table I are computed by making the simplest choice for the weights in the formula, that is $w_j = \frac{1}{(j^2 - j + 1)\alpha}$ for all j , where α is a normalization coefficient to make $S \in (0, 1)$. This choice has the effect of reversing the meaning of the evaluation scale: if S is near 0 network conditions can be assumed to be stationary; on the other side, larger values mean that radio conditions are getting worse. It is possible to set a threshold value on S to determine when channel performances are degraded.

The first column in Table I shows which BLER values are compared in the dispersion test. The other two columns report values obtained by applying (4) to compute the score, starting from p-values generated by the Moses and Miller test, respectively. It is clear from the results presented in Table I that the score S is a good indicator of performance degradation or improvement.

These results suggest that the Miller test is able to better discriminate different conditions; however the Moses test allows faster computation. It is important to highlight that a different choice of the window duration D or of the aggregation time interval T (for a given D) means a different size of sample vectors at each scale and so a different capability to detect variations or, roughly speaking, a different resolution.

For monitoring purposes it may be useful to set up an ob-

servation time span including a few consecutive temporal windows of observed traffic. This span should be moved forward if the S indicator does not reveal variations and extended whenever some sort of pre-alarm threshold is exceeded. In the latter case, trends can be better evidenced by multiple comparisons and an alarm can be raised when a meaningful change in network performances is confirmed. This situation can lead to more detailed troubleshooting analysis by means of complex instruments such as protocol analyzers.

TABLE I
MOSES AND MILLER TESTS SCORE COMPARISON

Block Error Rate [%]	S (Moses test)	S (Miller test)
0 vs 0	0.0279	0.0334
10 vs 0	0.4756	0.5006
30 vs 0	0.5066	0.5772
30 vs 10	0.1384	0.1685

V. FINAL REMARKS

The possibility to infer information about network performances from measurements at a few, or even a single monitoring point is, of course, of great interest to researchers as well as to network operators. In this respect, the mobile network architecture is well suited to the approach proposed in this paper: in fact, the open interfaces between the network elements are natural choices as monitoring points and represent different degrees of traffic aggregation. It should be remarked that observation of packet traffic at the TCP level is not always possible directly: in some cases sophisticated analyzers featuring multiprotocol analysis functions would be required to collect data in a suitable format for the application of a wavelet-based algorithm. However, when direct observation is possible, the resulting monitoring instrument becomes comparatively inexpensive, since the complexity of the analysis is essentially related to the algorithms, whereas processed data are simple and cheap to collect. Results presented in the paper are mainly based on an extensive set of simulations for an UMTS network, since very little real-world data were available at the time of this work. Care was taken to ensure that realistic traffic conditions were generated; in particular, the behavior of the relevant parts of the UMTS protocol stack was studied and parts of the ns-2 simulator were implemented from new, or modified, for the purpose of ensuring the meaningfulness of results. Future work will hopefully be able to rely on experimental packet traffic data to further validate the monitoring method discussed in this paper.

REFERENCES

- [1] P. Huang, A. Feldmann, W. Willinger "A non-intrusive wavelet based approach to detecting network performance." Proc. *ACM SIGCOMM Internet Measurement Workshop 2001*, San Francisco (CA), Nov. 2001.

- [2] K. W. Richardson, "UMTS review," *IEEE Electronics & Communication Engineering Journal*, Jun. 2000.
- [3] S. Mallat, "A wavelet tour on signal processing," Academic Press, 1998.
- [4] I. Daubechies, "Ten lectures on wavelets," SIAM Press, 1992.
- [5] P. Abry, D. Veitch, "Wavelet analysis of Long Range Dependent traffic," *IEEE Trans. on Information Theory*, Vol.44, n.1, Jan 1998, pp. 2-15.
- [6] P. Abry, P. Flandrin, M. S. Taquq, and D. Veitch, "Wavelets for the analysis, estimation and synthesis of scaling data," in *Self-Similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger Eds., New York: Wiley, 2000, pp. 39-88.
- [7] A. Feldmann, A. C. Gilbert, P. Huang and W. Willinger, "Dynamic of IP traffic: the role of variability and the impact of control," *Proc. ACM SIGCOMM 1999*, Cambridge (MA), Sept. 1999.
- [8] M. Hollander, D. Wolfe, "Non parametric statistical methods," 2 edition, Wiley & Sons, pp. 106-188, 2003.
- [9] , "Applied nonparametric statistics," Houghton Mifflin Company, pp. 69-128, 1978.
- [10] *Network Simulator 2* Home Page: <http://www.isi.edu/nsnam/ns/>.
- [11] *3GPP. Radio Link Control (RLC) Protocol Specification*, TS 25.322 v5.4.0 edition, 03 2003.
- [12] *3GPP. Medium Access Control (MAC) Protocol Specification*, TS 25.321 v2.2.0 edition, 10 1999.
- [13] A. Ahtianinen, L.Laitinen, S.Naghian, V.Niemi, "UMTS Networks: Architecture, Mobility and Services," John Wiley & Sons, 2001.