Multi-Sensor Fusion System Using Wavelet Based Detection Algorithm Applied to Network Monitoring

V Alarcon-Aquino and J A Barria

Communications and Signal Processing Research Group
Department of Electrical and Electronic Engineering
Imperial College, London, SW7 2BT UK
Email: v.alarcon, j.barria@ic.ac.uk

Abstract: A multi-sensor fusion system using wavelet based detection algorithm is proposed for network anomaly detection. The proposed approach is applied to monitor events in different network metrics of a Dial Internet Protocol service. The results show that the approach is able to identify the presence of abnormal behaviours in advance of reported network anomalies, and reduce the number of false alarms generated by each network metric.

1 Introduction

Proactive detection of network faults and performance degradations will enable service providers to take corrective action in advance of network/service disruptions. In this regard, several approaches have been proposed (see e.g., [1, 2]); however, these approaches are most suitable if the data contain contributions of events at fixed resolution or scale in time and/or frequency. Unfortunately, data from almost all practical network processes are multi-scale in nature, due to events occurring at different points in time and frequency. Recently, it has been demonstrated that a wavelet-based analysis seems more appropriate for data containing events whose behaviour changes over time and frequency (see e.g., [3]). In this paper, we have investigated a sensor fusion methodology applicable to network monitoring, which combines local decisions made from dispersed wavelet-based detectors (sensors) [3]. The time series are decomposed in time-frequency domain by undecimated discrete wavelet transform to capture localised transient events. Local decisions by wavelet-based sensors are followed by global decisions at the data fusion centre. The goal of the data fusion centre (DFC) is to improve system reliability by properly combining information from multiple wavelet-based sensors. Since the wavelet-based sensors are designed to work at different resolutions (in time and frequency), a better inference of events than any single sensor or combination of sensors with fixed resolution is obtained. The DFC incorporates the spatial dependencies between the monitored network metrics and provides temporally correlated alarms at different resolutions.

2 Proposed Approach

The proposed architecture for combining information at multiple resolutions from multiple wavelet-based sensors is illustrated in Figure 1. The S_i wavelet-based sensors ($i = 1, ..., n$), where n denotes the number of sensors, monitor different network metrics of a Dial Internet Protocol (IP) service. The global decision function or network health, $u_{0,j}(t) \in \{0, 1\}$, is obtained at different resolutions, $j = 1, ..., J$ where J represents the number of resolutions.

The output of the individual wavelet-based sensors, S_i, in time-correlated fashion, $u_{i,j}(t) \in \{0, 1\}, i = 1, ..., n$, are combined using the DFC. The DFC is composed of two parts. One is a weight (correlation) matrix in which the relationship between the monitored network metrics has been taken into account. The other part is the decision fusion rule, $\varphi(\cdot)$. Since identical local decisions, $u_{i,j}(t)$, and identically distributed observations after wavelet decomposition have been considered at all the wavelet-based sensors S_i, the optimal decision fusion rule, $\varphi(\cdot)$, reduces to k-out-of-n; that is, the network health $u_{0,j}(t) = 1$ if k or more wavelet-based sensor decisions are one [4].

The output of the individual wavelet-based sensors S_i, $u_{i,j}(t), i \neq 0$, is obtained by comparing the posterior probabilities after integrating with respect to a prior distribution. The posterior probability associated with the hypothesis H_0, which is computed using Bayes' theorem, can be written as follows:

$$p\left(H_0 | \left(W_{2^j} f_t\right)_i\right) \propto p\left(H_0\right) \int_0^\infty p\left(\left(W_{2^j} f_t\right)_i | t_0, \tau_j^2\right) p\left(\tau_j^2\right) d\tau_j^2, \quad t = 1, ..., N \text{ and } j = 1, ..., J \qquad (1)$$

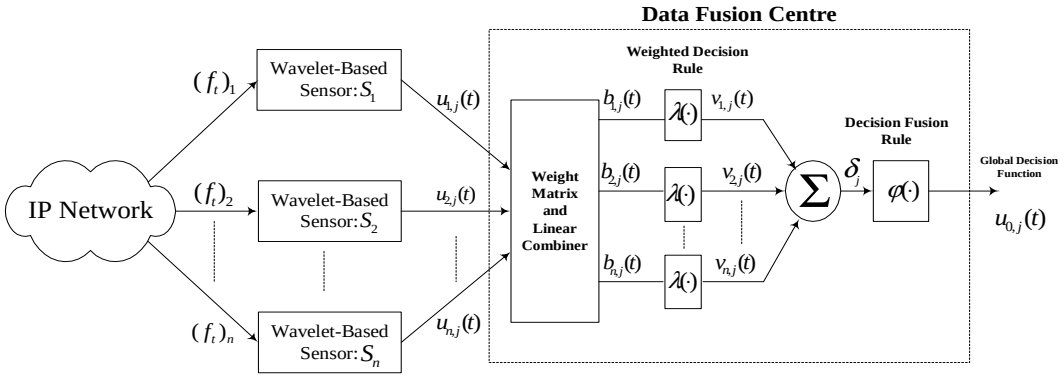

Figure 1: General architecture of multi-sensor data fusion system for network anomaly detection. $(f_t)_i$ denotes network measurements.

where $\propto$ denotes a relationship of proportionality, $p\left(H_0 \middle| \left(W_{2^j} f_t\right)_i\right)$ and $p\left(H_0\right)$ denote the posterior and prior probabilities associated with the hypothesis respectively, t_0 is an unknown change point, $p\left(\tau_j^2\right)$ is the prior distribution to be considered in the unknown variance of the wavelet coefficients,$\left(W_{2^j} f_t\right)_i$, and N denotes the length of the network measurements. The alternative hypothesis H_1 is obtained in a similar manner to Eq. (1) except that the unknown variance τ_j^2 of wavelet coefficients is changing at the time instant t_0, namely, $\tau_{j,a}^2$ for $t \leq t_0$ and $\tau_{j,b}^2$ for $t_0 < t \leq N$. The prior probabilities associated with the hypotheses are $p\left(H_0\right) = \pi_p$ for $\pi_p \in (0,1)$ and $p\left(H_1\right) = 1 - \pi_p$ with $p\left(H_0\right) + p\left(H_1\right) = 1$. Therefore, the prior probability of having a change point can be incorporated into the wavelet-based sensors. On the tests reported in this paper, the inverse Wishart distribution has been used as prior, and the quadratic spline wavelet has been used for the wavelet decomposition [5]. The posterior probabilities are then given by

$$p\left(H_0 \middle| \left(W_{2^j} f_t\right)_i\right) = \frac{p\left(H_0\right) S^{v/2} \pi^{-N/2} \Gamma\left(\left(N+v\right)/2\right)}{2^{(N+v)/2} \Gamma\left(v/2\right) \left[\left(\sum_{t=1}^{N} \left(W_{2^j} f_t\right)_i^2 + S\right) \Big/ 2\right]^{(N+v)/2}} \tag{2}$$

$$p\left(H_1 \middle| \left(W_{2^j} f_t\right)_i\right) = \frac{p\left(H_1\right) S^{v} \Gamma\left(\left(N-L+v\right)/2\right) \Gamma\left(\left(L+v\right)/2\right) \left[\left(\sum_{t=1}^{N-L} \left(W_{2^j} f_t\right)_i^2 + S\right) \Big/ 2\right]^{-(N-L+v)/2}}{2^{(N+2v)/2} \pi^{N/2} \Gamma\left(v/2\right) \Gamma\left(v/2\right) \left[\left(\sum_{t=N-L+1}^{N} \left(W_{2^j} f_t\right)_i^2 + S\right) \Big/ 2\right]^{(L+v)/2}} \tag{3}$$

where Γ denotes the gamma function, S and v represent the hyperparameters of the inverse Wishart distribution, and L is the length of a sliding window. Other priors and wavelets are assessed in [3]. The local decision rule is thus given by

$$\gamma_{i,j}\left(t\right) = \frac{\log p\left(H_0 | \left(W_{2^j} f_t\right)_i\right)}{\log p\left(H_1 | \left(W_{2^j} f_t\right)_i\right)} > 1, \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, J \tag{4}$$

which provides a basis for choosing between H_0 and H_1, and is bounded $u_{i,j}(t) = \gamma_{i,j}(t) \in \{0,1\}$. Each wavelet-based sensor employs a decision rule $\gamma_{i,j}(t)$ to make a decision $u_{i,j}(t)$, $i = 1, 2, \ldots, r$, where r denotes the number of wavelet-based sensors. Note that H_1 represents an abnormal behaviour in the network metric under consideration. The wavelet-based algorithm monitor events in a set of network metrics from a Dial IP service, namely, the *CT (Connect_Time)*, the *LT (Log_Time)*, the *DNST (Domain_Name_Server_Time)*, the *WL (Web_Latency)* and the *DT (Data_Time)*. Details of these network metrics can be found in [6]. The local decision vector is thus $\underline{u}_j(t) = [u_{1,j}(t)\ u_{2,j}(t)\ u_{3,j}(t)\ u_{4,j}(t)\ u_{5,j}(t)]^T$ where $u_{1,j}(t) = CT$, $u_{2,j}(t) = LT$, $u_{3,j}(t) = DNST$, $u_{4,j}(t) = WL$ and $u_{5,j}(t) = DT$.

The spatial dependencies between the monitored network metrics are computed using a correlation test and scatter diagram. The correlation test considered in this paper is the nonparametric Spearman's test [7]. The spatial dependencies between the monitored network metrics are taken into account by a weight (correlation) matrix, $W \in R^{n \times n}$. Strong correlations were observed between the network metrics *DNST*

and WL as well as between the network metrics WL and DT. This is to be expected because the WL metric involves the time taken for the first packed of data to return and it is strongly related to the time to download ($Data\_Time$). These network metrics ($DNST$, WL and DT) comprise the data transfer phase.

The output of the individual wavelet-based sensors, $u_{i,j}(t)$, are then weighted based on the spatial dependencies between the monitored network metrics. That is, using a linear combiner $\underline{b}_j(t) = W\underline{u}_j(t)$, where $\underline{b}_j(t) = [b_{1,j}(t) \quad b_{2,j}(t) \quad b_{3,j}(t) \quad b_{4,j}(t) \quad b_{5,j}(t)]^T$, the weighted decision rule, $v_{i,j}(t) = \lambda(b_{i,j}(t))$, is then expressed as

$$v_{i,j}(t) = \begin{cases} 1, & \text{if} \quad b_{i,j}(t) > 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where any time $b_{i,j}(t) > 0.5$, the network metric represents an abnormal condition. Thus, the decision fusion rule, $\varphi(\delta_j(t)) \in \{0, 1\}$, with $\delta_j(t) = \sum_{i=1}^{n} v_{i,j}(t)$, is given by

$$u_{0,j}(t) = \begin{cases} 1, & \text{if} \quad \delta_j\,(t) \geq 2k - n \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where $u_{0,j}(t)$ denotes the global decision function, and k denotes the number of wavelet-based sensors that decide hypothesis H_1 (see Eq. (4)). It is worth noting that this approach does not require setting any thresholds at the output of the individual wavelet-based sensors and therefore the problem is reduced to the choice of the value of k.

3 Application to Dial IP Service

Real world network data collected every 10 minutes over a period of 6 months from BT (British Telecommunications) Dial IP service are used in the following tests.[1] Figure 2 shows the output of the individual wavelet-based sensors (asterisks in figure) for the network metrics LT, WL and DT for the data set period from 23 October to November 05, 1999. The alarms of these network metrics are obtained by the local decision rule $\gamma_{i,j}(t)$ (see Eq. (4)).

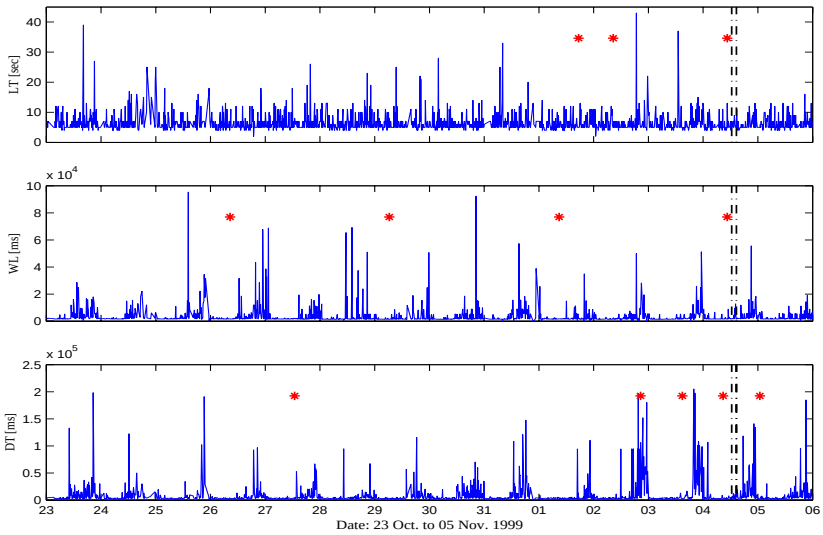

Figure 2: Output of the individual wavelet-based sensors for the network metrics LT, WL and DT. Vertical dash-dotted lines indicate abnormal period.

The global decision function $u_{0,j}(t)$ is shown in the upper plot of Figure 3 and is obtained once the spatial dependencies and temporal correlation have been taken into account. The lower part of Figure 3 shows the behaviour of the local decision rule for the monitored network metrics, $\gamma_{2,1}(t)$=LT, $\gamma_{4,1}(t)$=WL and $\gamma_{5,1}(t)$=DT. These results point out that for the Dial IP network metrics here investigated, the multi-sensor data fusion system is able to identify abnormal behaviours prior to the abnormal periods at

[1] It is worth remarking that BT quality of service cannot be inferred from the results reported in this paper.

resolution $j = 1$, and reduce the number of false alarms when compared to previous reported algorithm (see e.g., [3]). Similar results are obtained at resolution $j = 2$. An alarm is considered true alarm if this is within the interval of 60 minutes before and 25 minutes after the network anomaly. Otherwise, it is considered as a false alarm.

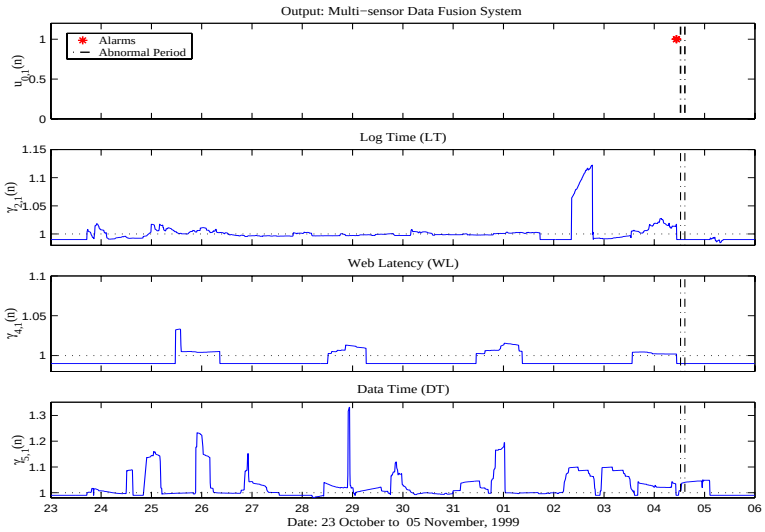

Figure 3: (Upper plot) Output of the multi-sensor data fusion system, $u_{0,j}(t)$, at resolution $j = 1$ using $k = 3$. Behaviour of the local decisions $\gamma_{2,j}(t)$, $\gamma_{4,j}(t)$, $\gamma_{5,j}(t)$ at resolution $j = 1$ for the network metrics LT, WL and DT respectively.

4 Conclusions

A multi-sensor data fusion system using wavelet-based detection algorithm has been proposed to reduce the number of false alarms generated by each network metric and to incorporate interdependencies between the monitored network metrics. The multi-sensor data fusion system provides temporally correlated alarms at different levels of resolution, and provides a better view and inference of the network health by looking at a global decision function.

Acknowledgements

The authors would like to thank I. Thurlow (BT) and G. Walker (BT Ignite, IP Service Platforms) for allowing these data sets to be used in this paper. The first author gratefully acknowledges the financial support from the National Council for Science and Technology (CONACYT), Mexico.

References

[1] Thottan, M. and Ji, C. Statistical Detection of Enterprise Network Problems. *Journal of Network and System Management*, 7(1):27-45, 1999.

[2] Ho, L. L., Cavuto, D. J., Papavassiliou, S. and Zawadzki, A. G. Adaptive and Automated Detection of Service Anomalies in Transaction-Oriented WAN's: Network Analysis, Algorithms, Implementation and Deployment. *IEEE Journal on Selected Areas in Communications,* 18(5):744-757, May 2000.

[3] Alarcon-Aquino, V. and Barria, J. Anomaly Detection in Communication Networks Using Wavelets. *IEE Proceedings - Communications*, 148(6):355-362, December 2001.

[4] Ramanarayanan, V. and Varshney, P. K. Distributed Detection with Multiple Sensors: Part I – Fundamentals. *Proceedings of the IEEE*, 85(1):54-63, January 1997.

[5] Mallat, S. and Zhong, S. Characterisation of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(7):710-732, 1992.

[6] Visual Internet Benchmark,http://www.visualnetworks.com/products/prod_inbench_internet.html

[7] Conover, W. J. *Practical nonparametric statistics*, Third Edition, Wiley Series in Probability and Statistics: Applied Probability and Statistics Section, John Wiley & Sons, Inc., 1999