# Real-time models supporting resource management decisions in highly variable systems

Sara Casolari, Michele Colajanni, Stefania Tosi
Department of Computer Engineering
University of Modena and Reggio Emilia
{sara.casolari,michele.colajanni,stefania.tosi}@unimore.it

Francesco Lo Presti
Department of Computer Engineering
University of Roma "Tor Vergata"
lopresti@info.uniroma2.it

## Abstract

*Data centers providing modern interactive applications are enriched by autonomous management decision systems that are able to clone and migrate virtual machines, to re-distribute resources or to re-map services in real-time. At the basis of all these decisions, there is the need of a continuous evaluation of the state of system resources and of detecting when some relevant changes are occurring. Unfortunately, the load of interactive applications reaching the system is intrinsically heterogeneous with consequent highly variable effects on the resource behavior emerging from system monitors. Hence, existing algorithms for online detection of state changes are affected by low precision and scarce robustness when they are applied to modern contexts. We propose a novel model for online detection of relevant state changes that combines a filtered representation of the raw measures with adaptive detection rules. Experiments carried out on real and emulated data sets confirm that the proposed model is able to timely signal all relevant state changes, to limit false detections and, even more important, its results are robust in highly variable contexts.*

## 1 Introduction

Any resource management mechanism related to anomaly detection, quality control, request redirection, diagnosis and fault detection, process migration, access control requires a continuous evaluation of the resource state and some real-time algorithm to detect as soon as possible whether some relevant change is occurring in the resource state.

We consider this problem in the specific context of modern multi-core architectures hosting several interactive Internet/Web-based applications on virtual machines. Here, we can assume that the most important system resources (CPU, memory, disk, network of guest and host machines) are continuously monitored and raw data sets are passed to some statistical algorithm that decides whether to signal to the resource management mechanism the occurrence of a relevant state change.

State change detection algorithms are affected by well known issues related to false positive and false negative detections. Online versions are also subject to temporal constraints. Moreover, in the considered context, the typical problems are exacerbated by raw data sets characterized by non-stationary and non-deterministic features, variable variance of the distribution, and likely affected by internal and external perturbations. The most common models for state change detection, such as Particle Filtering [2], Kalman [6] and Sequential Monte Carlo Method [16] are not effective in the considered context because they require an anticipated statistical analysis of the data set to set the right parameters of the equations [9]. Other state change detectors, such as the threshold-based detectors, the Shewhart chart, the Exponential Weighted Moving Average (EWMA) chart, and the baseline CUSUM are inadequate as well when data sets are highly variable. All of them tend to cause either a large number of false detections or absence of detections depending on the chosen parameters for the algorithm [3, 13]. In a pre-

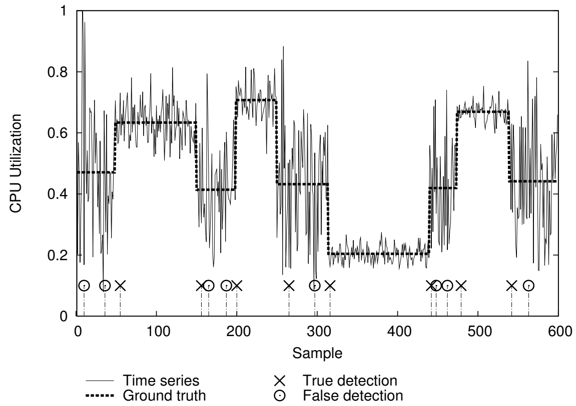



**Figure 1. The problem of state change detection**

vious work by the same authors [4], the proposal of an adaptive state change detector demonstrated the possibility of solving the tradeoff between excessive delay vs. false positive in non-stationary data sets. However, this adaptive solution does not work well when high variability and possible variable variance is combined with non-stationary effects in data sets.

Let us evidence the typical issues that affect state of the art detection algorithms through an example. We consider the CPU utilization measures of a server hosting five virtual machines. These measures are typically characterized by non-stationary behavior and by several intervals of high variability as shown by the spiked line in Figure 1. This figure contains also a line denoting the eight relevant state changes (at samples 50, 150, 200, 250, 315, 440, 475, 540) that the detection algorithm should be able to signal and that we consider as *ground truth*. We report the results of the online adaptive version [4] of the detection CUSUM algorithm demonstrated to be optimal in stationary conditions [15]. The small vertical line with a circle at the top denotes a false detection, that is, a signaled change that does not correspond to a real state change. The vertical line with a cross at the top denotes a right detection. Figure 1 evidences the two main problems that may affect an online state change detection algorithm when the data set is non-stationary and highly variable: several false detections (8 lines with a circle) and high detection delays. For example, the algorithm correctly signals a state change at sample 265 when it actually occurred at sample 250, and 15 samples of delay may be unacceptable in many contexts. Addressing these issues requires a novel model that pursues two objectives:

- an *online filter* that removes perturbations from the raw data set and produces a smoothed data representation;
- an *adaptive detection model* that is able to modify dynamically its parameters.

The real-time detection model we propose in this paper combines an online Wavelet-based data representation for the filtering phase and an adaptive implementation of the CUSUM for the detection phase. We demonstrate that the proposed Wavelet-based adaptive model is able to guarantee the best results for several statistical characteristics of the time series in simulated and real contexts, and it satisfies the temporal constraints imposed by the real-time management mechanisms.

The rest of the paper is organized as follows. Section 2 describes the proposed detection model. Section 3 presents the performance metrics and results of the proposed model against state of the art algorithms applied to several sets of emulated data. Section 4 analyzes the results of the proposed model for real data sets. We conclude the paper with some final remarks.

## 2 State change detection model

Many resource management mechanisms are integrated with algorithms that are able to decide whether some relevant change has occurred in the profile of some monitored system resources. We are interested to these algorithms applied to non-stationary and highly variable data sets, where each data set is a time series that at sample $i$ can be denoted by $\{y_i\} = (y_1, \ldots, y_i)$. In the considered contexts, it is impossible to identify a stable state if not for short periods [1]. Hence, the *state* of a resource is a reference representation (e.g., mean, variance, confidence interval) that is valid for the specific interval during which the statistical attributes of the time series do not change significantly. Hence, we can see the time series as a sequence of consecutive stable states, where a *relevant state change* corresponds to a significant variation of the statistical characteristics of the time series [3].

The performance of state change detection algorithms is highly dependent on the statistical characteristics of the data set [13]. Because of the inherent variability and non-stationary behavior of the monitored processes related to interactive systems, existing algorithms achieve poor results. We propose a new algorithm, namely *Wavelet-Adaptive CUSUM*, that integrates and extends two well known models (Wavelet [12] and CUSUM [15]) in an adapted version. This algorithm filters raw data through an online Wavelet model (Section 2.1). The denoised representation of the raw measures is then passed to an adaptive and online implementation of the CUSUM detection rule that must signal just relevant state changes as soon as possible (Section 2.2).

## 2.1 Filtering model

We claim that non-stationary and highly variable data sets require online statistical techniques that are able to distinguish and isolate all the perturbations from the main features of the time series. The literature presents dozens of data filters, but few of them can be suitably used in real-time contexts. For example, Exponential Weighted Moving Average (EWMA) models are applied in several online contexts, such as in information and computer systems, and financial and social applications [1, 11]. They are appreciated for their simplicity and computational efficiency, but they correspond to single scale low pass filters. As a consequence, if a time series contains features at multiple scales, EWMA filters must tradeoff the extent of noise removal with the quality of the retained features [14]. In a typical context such as that shown in Figure 1, this would either results in reduced noise removal and too many false detections or in excessive smoothing which would cause too many negative detections.

Non-linear and multiscale models exhibit better performance than the linear counterparts. In this paper, we are interested to the Wavelet-based models that have been shown to be nearly optimal for various error norms and smoothness of the resulting time series [8]. The main motivation for this choice is that Wavelet denoisying exploits an orthonormal basis localized both in space and frequency which allows a noise reduction without smoothing the time series features. On the other hand, the solutions based on exponential smoothing work only in the frequency domain. As a consequence, we do not pass to the state change detector the original data set $\{y_i\}$, but a *filtered* version $\{x_i\}$ [14], where $\{x_i\}$ retains the significant features of the original data but it removes most of the variability that can be ascribed to short-term perturbations. We regard $\{x_i\}$ as the *data representation* of the monitored process obtained through an online version of the Wavelet transformation [14]. It represents a time series as the sum of a shifted and scaled version of a base Wavelet function $\psi$ and a shifted version of a low-pass scale function $\phi$. Through a proper choice of the Wavelet and scale functions, the resulting families of functions are:

$$\psi_{mk}(i) = \sqrt{2^{-m}}\psi(2^{-m}i - k) \qquad (1)$$

$$\phi_{mk}(i) = \sqrt{2^{-m}}\phi(2^{-m}i - k) \qquad (2)$$

where $m$ and $k$ are the dilation and translation parameters from an orthonormal basis, respectively. Hence, the time series $\{y_i\}$ can be conveniently rewritten as follows:

$$y_i = \sum_{k=1}^{2^{-L}n} a_{Lk}\phi_{Lk}(i) + \sum_{m=1}^{L}\sum_{k=1}^{2^{-m}n} d_{mk}\psi_{mk}(i) \qquad (3)$$

where $a_{Lk}$ is the $k$-th scaling function coefficient at the coarsest scale $L$, $d_{mk}$ is the $k$-th Wavelet coefficient at scale $k$, and $n$ is the length of the time series considered for the analysis. We set the coarsest scale $L = 4$ as suggested in [14]. The coefficients $m$ and $k$ are computed by the inner product of $\{y_i\}$ with the base functions. Computation of the transform and its inverse can be done in $O(n)$. A key feature of this representation is that the Wavelet decomposition captures significant signal features in a few relatively large coefficients, while perturbations result uncorrelated. As a result, perturbations - and perturbations only - can be effectively removed by setting equal to zero the Wavelet coefficients smaller than a threshold.

In summary, we obtain the data representation $\{x_i\}$ of the original time series $\{y_i\}$ through three main steps. We apply to $\{y_i\}$ the standard Haar function [5] as a base Wavelet, which consists of a simple rectangular impulse function. Then, we set to zero the Wavelet coefficients which are lower than a suitable threshold $t_m$, where $m$ is the dilation parameter. As indicated in [14], we set the threshold $t_m = \sigma_m\sqrt{2\log n}$ where $\sigma_m = \frac{1}{0.6745}\text{median}\{|d_{mk}|\}$. Finally, we compute the inverse Wavelet transform to obtain $\{x_i\}$.

This filtering technique has been proven to be superior to other approaches [8], but in literature it is restricted to offline operations. We consider an online version which takes into account a moving window of dyadic length to estimate the Wavelet parameters and to compute $\{x_i\}$. At each sample $i$, the filtered value $x_i$ is computed as follows:

1. consider the set $\{y_i\} = (y_{i-M+1}, \ldots, y_i)$ of maximum dyadic length, where $M = \lfloor \log_2 i \rfloor$;
2. compute the filtered sequence $(z_1, \ldots, z_M)$ of the set $\{y_i\}$ through the three previous steps;
3. set $x_i = z_M$, that is, set the actual filtered value equal to the last value of the filtered sequence computed at step 2.

This online version can be computed in $O(n \log n)$ steps (see [14] for details), hence it is suitable to temporal constraints required by real-time management systems.

## 2.2 State change detection

There are many state change detection algorithms proposed in literature. Among them, we refer to the well known CUSUM model that has been demonstrated to achieve optimal results in specific conditions of the data set [15]. As most of previous algorithms, the baseline version of CUSUM works offline and, as a consequence, we can set its best parameters on the basis of the statistical features of the data set [3]. We proposed an online version of the CUSUM model (Adaptive CUSUM) that was able to compute dynamically its parameters [4]. That model improved the previous results but preliminary experiments (not reported for space reasons) demonstrated that it does not work well when the data set is highly variable. For these reason, we propose to apply the adaptive version of the CUSUM not to the raw data set but to the data representation emerging from the application of the online Wavelet-based filter. We name *Wavelet-Adaptive CUSUM* the combination between the Adaptive CUSUM and the online Wavelet filter. If we consider that a state at sample $i$ is represented by its mean value $\mu_i$ (continuously estimated online, as suggested in [4]), the adaptive implementation of CUSUM for detecting an increase in the mean of the Wavelet-based representation $\{x_i\}$ uses the following test statistics (the case for detecting a decrease is dual):

$$g_0^+ = 0 \quad (4)$$

$$g_i^+ = \max\{0, g_{i-1}^+ + x_i - (\mu_i + K)\} \quad (5)$$

The gain function $g_i^+$ accumulates deviations of the Wavelet representation $x_i$ from the online estimation of the state value $\mu_i$ that are greater than a predefined threshold $K$, and resets to 0 on becoming negative. The term $K$, which is known as the allowance or slack value, determines the minimum deviation that the statistic $g_i^+$ accounts for. A positive change is signaled when $g_i^+$ exceeds a threshold $H$. To achieve a good detection quality, the suggested value for $K$ is $\frac{\Delta}{2}$, where $\Delta$ is the minimum shift to be detected. The choice for $H$ is explained below. A two-sided test to detect increases and decreases is obtained by applying the tests in 5 for increase and its dual version for decreases.

In the considered environment, that is characterized by non-stationary and highly variable time series, it is important to adopt an adaptive version of CUSUM to be combined with the online Wavelet model. This requires two interventions: a dynamic estimation of the new state representation, and a dynamic evaluation of the $H$ parameter.

The reference value of a stable state can be any statistical representation of the corresponding data set. In this paper, we consider the mean value $\mu_i$ as the reference value. Hence, we find convenient to refer to the Exponential Weighted Moving Average (EWMA) for the dynamic estimation of the new state representation because this model is able to track the slow varying mean even in the presence of non-stationary monitored samples. As soon as a relevant state change is signaled, we estimate online the reference value of the new state $\mu_i$ as follows:

$$\mu_i = \begin{cases} \mu_{i-1} + K + \frac{g_i^+}{N^+} & \text{if } g_i^+ > H \\ \mu_{i-1} - K - \frac{g_i^-}{N^-} & \text{if } g_i^- > H \end{cases} \quad (6)$$

where $N^+$ ($N^-$) denotes the number of steps elapsed since the last time $g_i^+$ ($g_i^-$) was set to zero, that is $N^+ = i - \inf\{j \mid g_j^+ = 0\}$ and similarly for $N^-$.

The choice of the $H$ value is crucial for the performance of the algorithm because the number of false detections directly depends on $H$ and on the standard deviation $\sigma$ of the time series. In the basic CUSUM version, the suggested value for $H$ is estimated on the basis of the Kullback approximation [3] that returns $H = 5\sigma$, where $\sigma$ is an estimation of the standard deviation during a previous stable interval of the time series [15].

On the other hand, when the standard deviation may change and when the time series perturbations reach high intensity (often equal or higher than the state change magnitude), any fixed choice for $H$ is inappropriate. We propose an adaptive evaluation that is based on the Sigmund approximation explicitly devoted to manage time series characterized by high perturbations [3]. In this equation, we introduce a continuous online estimation of the standard deviation at sample $i$ as in [10] that is, $\sigma_i = E[\mid \{y_i\} - E[\{y_i\}] \mid]$. It is computed on the same sliding window of data that is used for the online estimation of $\mu_i$. More details are reported in [4].

# 3 Performance comparison

In this section, we present some existing detection algorithms for comparison purposes (Section 3.1) and the main performance metrics (Section 3.2) computed for their evaluation. Then, we evaluate the performance of the state change detection algorithms on a wide range of data sets emulating non-stationary and highly variable profiles related to system resources (Section 3.3).

### 3.1 Other algorithms

We outline other popular state change detection algorithms based on CUSUM that can be applied to online data sets and that we consider for comparison purposes in our analysis.

The **Baseline CUSUM** applies the CUSUM detection rule on raw data [13, 15]. Its detection quality is limited by the high variability of the data set and by the static choice of parameters.

The **Adaptive CUSUM** adds an adaptive choice of the CUSUM parameters [4] but the detection rule works on monitored samples. Its performance are limited by the absence of a data filter.

The first idea is to add to the previous algorithm an online filter that is based on an Exponential Weighted Moving Average (EWMA) computed on the last $n$ samples [13]. We denote the data representation as $\text{EWMA}_n$ and the overall algorithm as $\mathbf{EWMA}_n$**-Adaptive CUSUM**. The performance of this algorithm are limited by the linearity of the data representation in highly variable contexts. Moreover, it is too much sensitive to the choice of $n$. Using a small set of $n$ past values offers a more reactive data representation that tends to minimize the detection delay but at the cost of a high number of false detections. Increasing the number of $n$, the data representation gets more smoothed; this causes a lower number of false detections but higher detection delays and more false negatives. In the considered context where data sets are highly variable and non-stationary, it is impossible to find a right value for $n$ that is able to guarantee reliable and efficient results. In this paper, we consider two different models: $\text{EWMA}_5$-Adaptive CUSUM should limit the delay for detection at the cost of a high number of false detections; $\text{EWMA}_{30}$-Adaptive CUSUM should provide a low or null number of false detections, although it may cause high detection delays.

### 3.2 Performance metrics

We evaluate the detection quality of all the algorithms in terms of *mean delay for detection* and *percentage of false detections* [3]. We are not interested to report results about false negative detections because in the time series used in this paper all considered algorithms are not affected by this problem.

The mean delay for detection, $T$, is related to the ability of an algorithm to detect a state change when it actually occurs. It quantifies the time delay for the detection of a new state through the distance between the sample at which the model signals a state change and the actual sample of change in the ground truth and computes the mean over all the state changes. For example, let us consider a time series with $X$ state changes. Let $[c_1, \ldots, c_X]$ be the actual samples of change in the ground truth and $[c_1^*, \ldots, c_X^*]$ the samples at which the model detects the changes. The mean delay for detection is defined as:

$$T = \frac{\sum_{i=1}^{X}(c_i^* - c_i)}{X} \tag{7}$$

Good detection algorithms should minimize $T$.

The percentage of false detections $P$ is obtained by the ratio between the number of false detections and the total number of detections signaled by the state change detector. The changes that are signaled correctly by the detection algorithm are called *true positives* ($TP$), otherwise they are classified as *false positives* ($FP$). Hence, $P$ can be defined as:

$$P = \frac{FP}{TP + FP} * 100 \tag{8}$$

where $(TP + FP)$ is the total number of detections. A value of $P$ equal to 0 means that the algorithm detects only relevant state changes, while high values of $P$ indicate that the detection algorithm signals many irrelevant state changes.

The best algorithm for state change detection should minimize the detection delay and the percentage of false detections. As there is a well known tradeoff between these two measures and the minimization of both of them is impossible, the best algorithm should find an acceptable compromise.

### 3.3 Comparative results

We compare the detection quality of the considered algorithms for a wide range of data sets emulating non-stationary and highly variable profiles related to system resources. To facilitate the algorithm comparison, the profile is normalized so that state increases/decreases are denoted by a unit value. All detection algorithms tend to diminish their detection quality for increasing variability of the time series. Hence, it is important to apply the algorithms on time series characterized by different levels of variability. The most important statistical properties that characterize a time series are the dispersion $\sigma$ and the correlation $\rho$ of the noise component [7]. To demonstrate the importance of using a filtered data representation to improve the detection quality, we evaluate the performance of the detection algorithms as a function of several values of $\sigma$ and $\rho$, but our main interest is on time series characterized by high dispersion and correlation of the noise component, for example $\sigma > 0.5$ and $\rho > 0.1$.

| $\sigma$ | Baseline CUSUM | Adaptive CUSUM | EWMA$_5$ Adaptive CUSUM | EWMA$_{30}$ Adaptive CUSUM | Wavelet Adaptive CUSUM |
|---|---|---|---|---|---|
| **0.1** | 44 | 39 | 0 | 0 | 0 |
| **0.2** | 52 | 40 | 0 | 0 | 0 |
| **0.3** | 63 | 42 | 2 | 0 | 0 |
| **0.4** | 75 | 57 | 6 | 0 | 0 |
| **0.5** | 79 | 64 | 15 | 0 | 0 |
| **0.6** | 89 | 75 | 33 | 0 | 2 |
| **0.7** | 93 | 78 | 49 | 0 | 8 |
| **0.8** | 100 | 80 | 55 | 0 | 8 |
| **0.9** | 100 | 85 | 69 | 0 | 13 |
| **1.0** | 100 | 97 | 74 | 0 | 19 |

**Table 1. Percentage (%) of false detections -** $\rho = 0.2$

| $\sigma$ | Baseline CUSUM | Adaptive CUSUM | EWMA$_5$ Adaptive CUSUM | EWMA$_{30}$ Adaptive CUSUM | Wavelet Adaptive CUSUM |
|---|---|---|---|---|---|
| **0.1** | 1.29 | 1.19 | 6.17 | 27.75 | 5.35 |
| **0.2** | 3.15 | 3.10 | 6.78 | 28.15 | 8.62 |
| **0.3** | 4.12 | 4.61 | 8.03 | 30.52 | 10.75 |
| **0.4** | 4.89 | 5.55 | 8.43 | 30.57 | 11.48 |
| **0.5** | 4.86 | 6.12 | 9.42 | 31.97 | 13.10 |
| **0.6** | 6.65 | 7.75 | 10.18 | 33.82 | 14.40 |
| **0.7** | 6.75 | 7.87 | 10.98 | 35.32 | 15.82 |
| **0.8** | 7.00 | 9.00 | 13.80 | 40.80 | 16.71 |
| **0.9** | 10.00 | 12.33 | 14.26 | 44.90 | 17.28 |
| **1.0** | 11.50 | 13.87 | 14.51 | 45.73 | 18.12 |

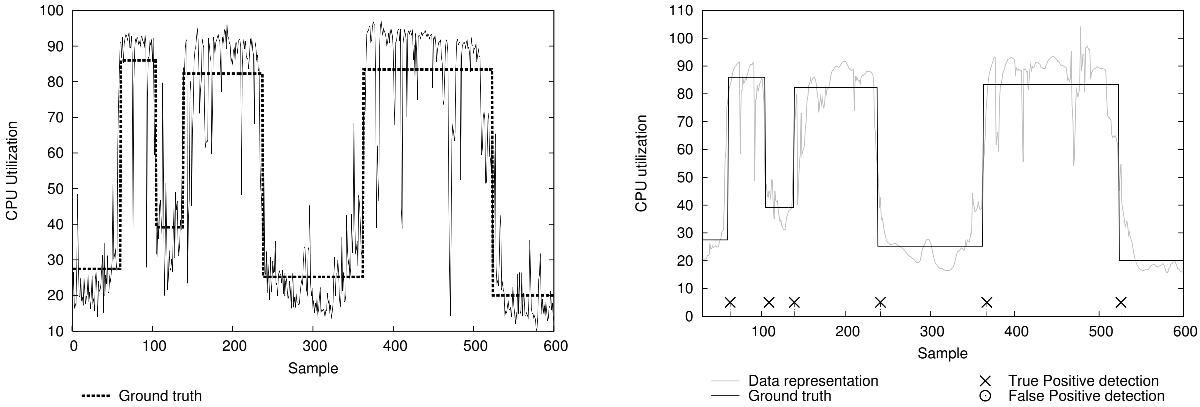
**Table 2. Mean delay for detection -** $\rho = 0.2$

In Table 1 we report the percentage of false detections when the correlation of the noise component $\rho$ is equal to 0.2. This table demonstrates the expected trait of all the detection algorithms: at higher noise dispersions, the percentage of false detections increases as well. However, the results are quite different. The models working on raw data sets (Baseline CUSUM and Adaptive CUSUM) reach a percentage of false detections always higher than 0.4 even for low $\sigma$, and they become useless for highly variable contexts. Similar unreliability is shown by the EWMA$_5$-Adaptive CUSUM.

The introduction of a filtered data representation impacts positively on the percentage of false detections. In particular, EWMA$_{30}$-Adaptive CUSUM and Wavelet-Adaptive CUSUM are able to achieve good performance in time series characterized by a noise dispersion $\sigma \leq 0.5$ and high correlation of noise. Surprisingly, EWMA$_{30}$-Adaptive CUSUM that uses a more smoothed data representation, is able to avoid false detections. However, this result comes at a price because we know that a trade-off exists. We have to consider also the mean delay results, that are reported in Table 2 for all the considered algorithms. The EWMA$_{30}$-Adaptive CUSUM is affected by a mean delay of 45 samples in the worst case. This result shows that this algorithm is inadequate to support real-time management decisions. Moreover, it confirms that the choice of the filtering technique is crucial for online state detection models. In highly variable contexts, just the proposed Wavelet-Adaptive CUSUM guarantees an efficient tradeoff between a low percentage of false detections and minimum delay. In the next section we aim to confirm these positive results by applying the proposed algorithm to data sets coming from monitors of real system resources.

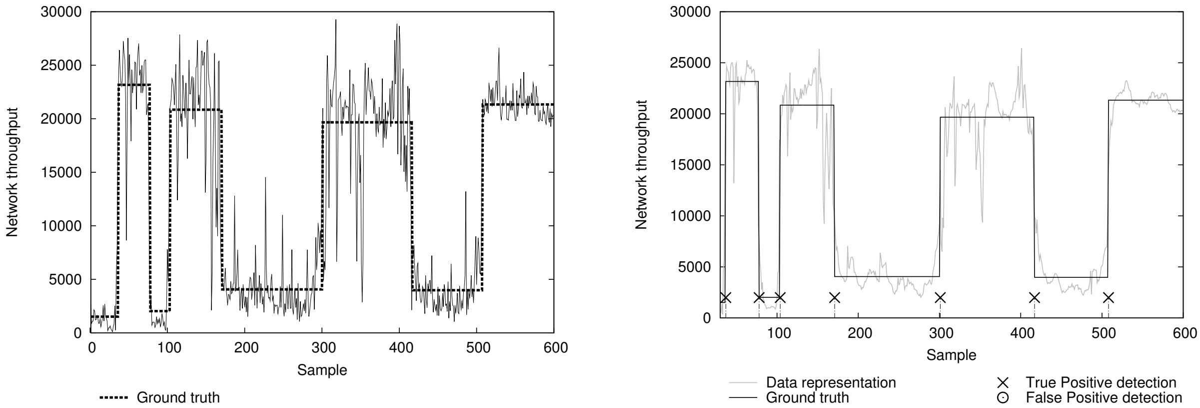
## 4 Experimental results

We applied the proposed algorithm to the resource measures of an Internet data center hosting Web-based applications. Here we report a subset of significant results. Figure 2(a), Figure 3(a) and Figure 4(a) show the data sets related to CPU utilization, network throughput and system call of a server hosting multiple guest servers, respectively. In these figures, the horizontal/vertical dotted line represents the ground truth, that is, the sequence of relevant stable states that the state change detectors should be able to signal. (In particular, the significant state changes occur at samples $[61, 105, 139, 238, 363, 524]$ in Figure 2(a), at samples $[37, 78, 104, 171, 301, 417, 508]$ in Figure 3(a) and at samples $[25, 59, 98, 139, 164, 211, 332, 488]$ in Figure 4(a)). The results of the proposed Wavelet-Adaptive CUSUM algorithm are reported in the correspondent figures (b). These figures contain two types



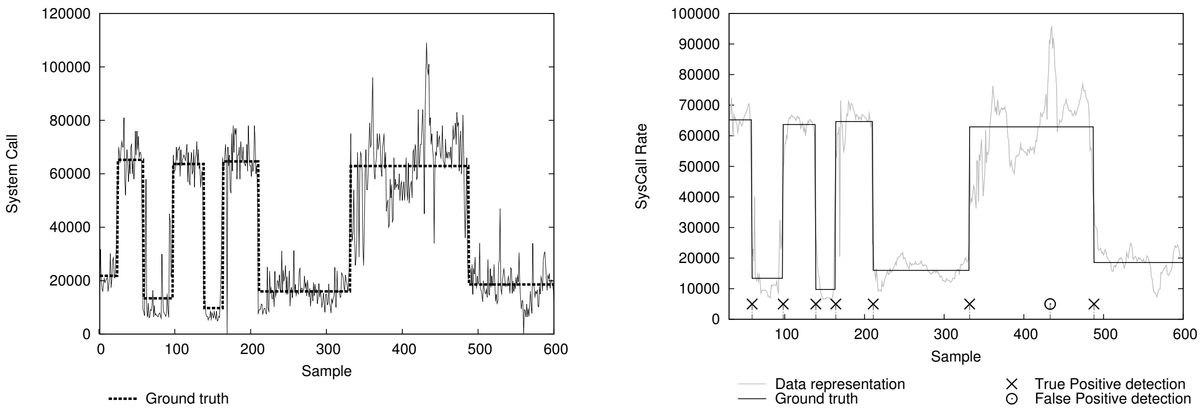

**(a) Resource measures** **(b) State change detections**

**Figure 2. CPU utilization**




**(a) Resource measures** **(b) State chage detections**

**Figure 3. Network throughput**




**(a) Resource measures** **(b) State change detections**

**Figure 4. System call**

of information: they report the data representation obtained through the online Wavelet filter and the state change detections that may be true or false.

It is interesting to observe that the Wavelet model applied to any data set is able to adapt the data representation to the system state conditions and to filter out the main data perturbations. In all the reported instances, the proposed algorithm detects timely all the state changes (there is no significant delay). Even more important, there is almost no false detection: indeed, just one at sample 433 in the system call case shown in Figure 4(b). This false detection is caused by a huge spike in the raw data that even the Wavelet model is unable to filter. We can add that, in the same conditions of the considered data sets, no other state change detection algorithm is able to achieve similar results because of excessive delay or excessive number of false signals.

The ability of guaranteeing good detection quality in extremely different and highly variable contexts demonstrates the robustness of the Wavelet-Adaptive CUSUM and its ability to provide a reliable support to real-time resource management decisions.

## 5 Conclusions

Modern data centers are equipped with real-time management decisions relying on continuous monitors of the system resources and detection algorithms that evaluate the resource states and relevant changes. The problem is that hosting multiple interactive applications causes system resource profiles characterized by non-stationary and highly variable behavior. We have proposed a new state change detection algorithm that is specifically tailored to work online in noisy contexts and that is characterized by low computational complexity. The proposed algorithm integrates an online version of the Wavelet model to filter the measures flowing from the system monitors and an adaptive version of the CUSUM statistical test as a state change detector. All experiments carried out on emulated and real data sets demonstrate that the proposed solution is robust and effective. It signals just relevant state changes and it is not affected by false negative detections. Moreover, it guarantees an efficient solution to the tradeoff between the percentage of false detections and the detection delay that previous algorithms do not address well.

## Acknowledgements

The authors acknowledge the support of MIUR-PRIN project DOTS-LCCI "Dependable Off-The-Shelf based middleware systems for Large-scale Complex Critical Infrastructures".

## References

[1] M. Andreolini, S. Casolari, and M. Colajanni. Models and framework for supporting run-time decisions in web-based systems. *ACM Tran. on the Web*, 2(3), 2008.

[2] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2), 2002.

[3] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes:Theory and Application*. Prentice-Hall, 1993.

[4] S. Casolari, M. Colajanni, and F. Lo Presti. Runtime state change detector of computer system resources under non stationary conditions. In *Proc. of 17th Int. Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, Sept. 2009.

[5] C. K. Chui. *An Introduction to Wavelets*. Academic Press, 1992.

[6] C. K. Chui and G. Chen. *Kalman filtering with real-time applications*. Springer-Verlag New York, Inc., 1987.

[7] M. Dobber, R. Van det Mei, and G. Koole. A prediction method for job runtimes in shared processors: Survey, statistical analysis and new avenues. *Performance Evaluation*, 2007.

[8] D. L. Donovo, I. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: Asymptotia? *Journal of the Royal Statistical Society B*, 57(2), 1995.

[9] F. Gustafsson. *Adaptive Filtering and Change Detection*. John Wiley and Sons, 2000.

[10] V. Jaconson. Congestion avoidance and control. In *Proc. of SIGCOMM'88*, volume 21, Stanford, CA, Aug. 1988.

[11] M. Kendall and J. Ord. *Time Series*. Oxford University Press, 1990.

[12] S. G. Mallat. A theory of multiresolution signal decomposition: The wavelet decomposition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7), 1989.

[13] D. C. Montgomery. *Introduction to Statistical Quality Control*. John Wiley and Sons, 2008.

[14] M. N. Nounou and B. Bakshi. On-line multiscale filtering of random and gross errors without process models. *American Institute of Chemical Engineers Journal*, 45(5), May 1999.

[15] E. S. Page. Estimating the point of change in a continuous process. *Biometrika*, 44, 1957.

[16] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo Method*. New York: John Wiley & Sons, 2007.