Context Information Management for Pervasive Computing

Odysseas Sekkas*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
sekkas@di.uoa.gr

Abstract. Pervasive Computing systems have to deal with the contextual information (context), which characterizes the current situation of the involved entities (e.g., users, mobile devices, environment, etc.). This dissertation studies context management issues related to the capability of a pervasive system on adapting its behavior to the involved entities context. Specifically, the interaction between the user and such system has to be less intruding as long as the latter recognizes the current user situation and adapts its functions accordingly. Such issues comprise the concept of Context Awareness. The dissertation focuses on context knowledge representation and reasoning as well as on approximate reasoning (Fuzzy Sets Theory). The management of lower level environmental information that emanates from sensors is also of great importance and is achieved with novel data fusion and decision fusion techniques that are proposed. Moreover, have been studied issues regarding collaborative context awareness and reasoning as well as bio-mimetic contextual dissemination. Adaptive algorithms are proposed so that is rendered possible the efficient dissemination of information in distributed environments. The objective is the minimization of communicating costs and the enhancement of context quality. Consequently, have been designated issues such as context discovery, context representation and inference, context fusion and collaborative context awareness

Keywords: pervasive computing, context management, context-awareness, context reasoning, epidemic algorithms

1 Introduction

In recent years we have witnessed a rapid progress in the pervasive (ubiquitous) computing paradigm. Specifically, pervasive computing is emerging as the future computing paradigm in which infrastructure and services are seamlessly available anywhere and anytime, improving human quality of life transparently to the underlying technologies. A system that is unobtrusively embedded in the environment, intuitive, constantly available and realizes the so-called ambient intelligence is defined as a pervasive system. The most profound technologies

* Dissertation Advisor: Stathes Hadjiefthymiades, Assistant Professor

are those that disappear [1].This exciting paradigm steps from an amalgamation of information and communication technology. It is not only the consequence of convergence of advanced electronics but is also the result of contemporary research and technological advances in wireless and sensor networks, distributed systems, mobile and agent computing, autonomic and context-aware computing.

Context-awareness is one of the basic factors of the pervasive computing. It is defined as the ability of a system to use any piece of information (context) by sensing the physical environment and adapt accordingly its behavior. In order to engineer context-aware systems, it is highly important to understand and define the ingredients of context from an engineering perspective. Context defines ambient conditions and describes the situation of an entity [2]. Contextual information might change over time, describing human behaviors, application and environmental states. Context fusion is the method of deducing new and relevant information from a variety of sources in order to be used by applications and users.

In this dissertation we studied context-related issues regarding the current situation of a user (e.g. location, actions). In such case is proposed a system which exploits data streams derived from sensors, in order to accurately estimate the location of a user. The term sensors includes Wi-Fi adapters, IR receivers, RFID tag readers, etc. The core of the system is the fusion engine which is based on Dynamic Bayesian Networks (DBNs), a powerful mathematical tool for integrating heterogeneous sensor observations [3]. An extension to this system is novel context fusion engine that models, determines and reasons about the user situation. This engine which is based on Dynamic Bayesian Networks and Fuzzy Logic, deals with the reliability of sources and approximate contextual reasoning [4],[5].

For the ambient context awareness is proposed a two-level fusion scheme. To cope with heterogeneous sensors (e.g. temperature, humidity) and deliver alarms with increased accuracy and confidence, a layered fusion scheme has been adopted [6]. Different sensor feeds are processed in the two layers of the fusion scheme thus improving the reliability of the system in detecting various events. On the lower layer, the statistical behavior of sensor data is constantly assessed. On the higher layer, Dempster-Shafer (DS) theory of evidence is adopted in order to mix the indications coming from the lower layer. The proposed system has been tested for fire detection [7],[8].

Moreover, have been studied issues regarding collaborative context awareness and reasoning as well as bio-mimetic contextual dissemination. Adaptive algorithms are proposed so that is rendered possible the efficient dissemination of information in distributed environments. The objective is the minimization of communicating costs and the enhancement of context quality. Consequently, have been designated issues such as collaborative context awareness [9],[10].

The following sections describe analytically an event detection schema for context awareness. Specifically, we adopt the CUSUM test for change detection in sensor data. An improvement of this technique is also proposed. Fusing the retrieved data data from neighboring sensors we are able to mitigate problems that lead to missed events and false alarms. Simulation results reveal the appropriateness of the mechanism in order to detect an event (particularly fire) as soon as possible and with low false alarm rate.

2 Data Fusion for Context Awareness

2.1 Event Detection

Let $\{X_i\}$ denote a sequence of random variables, i.e., a sequence of independent measurements of a sensor. We assume that X_i have density $f(x_i; \mu_0, \sigma)$ for $i = 1, \ldots, \tau - 1$ and density $f(x_i; \mu_F, \sigma)$ for $i \geq \tau$, where parameter μ_0 is known and μ_F and σ are generally unknown. The time index τ signals the event(e.g. fire) in which a change in the distribution of the measurements occurs.

The parameter, μ_0 may denote the mean data value which is estimated every T_0 sec (i.e., $T_0 = 30$ min) based on sensor measurements. This time window length is in general variable and it is advisable to decrease it during daily periods that are characterized by large variations (i.e. temperature from 5:00am to 12:00am). The parameter μ_F denotes the mean value in case of event and it is considered unknown. Similarly, σ^2 denotes the unknown variance of the measurements. For example, the Sensirion SHT11 temperature sensor has an accuracy of $\pm 2.5^o$C in the range from $-40^o to - 120^o$C. Adding a margin of 3^oC to accommodate variations due to clouds etc., we may assume that $\sigma = 5.5^o$. Nevertheless, σ is a nuisance parameter and it is generally unknown. If an event occurs then the parameter τ is the time index indicating a change of densities. Sequential tests can deal with this detection of change as discussed below.

One of the most promising algorithms to sequentially detect the change is the CUSUM test [11]. For instance, if the parameter of interest is the mean value, we can monitor the partial sums

$$S_n - \min_{1 \leq k \leq n} S_k, \qquad n = 1, 2, \ldots$$

where $S_n = \sum_{i=1}^{n} X_i$ and conclude that a change from the initial μ_0 mean value to μ_F occurs at time n (as long as the previous statistic is large enough).

Gombay [12] adapted Page's CUSUM test ([11]) for change detection in the presence of nuisance parameters. Gombay proposed statistics based on the efficient score (Rao's statistics), on the maximum likelihood estimator (Wald's statistics), or on the log likelihood ratio. The efficient score vector is defined as

$$V_k(\mu, \sigma) = \sum_{i=1}^{k} \nabla_\nu \log f(X_i; \mu, \sigma), \qquad \nu = (\mu, \sigma) \tag{1}$$

As it can be proved, if the density $f(\cdot)$ belongs to the exponential family, i.e., Gaussian, then once some regularity conditions hold under the null hypothesis, there exists a Wiener process $W(t)$ that approximates

$$W_k = \Gamma^{-1/2}(\mu_0, \sigma) V_k(\mu_0, \hat{\sigma}_k) \tag{2}$$

where $\hat{\sigma}_k$ is the maximum likelihood estimation of σ and $\Gamma(\mu_0, \sigma)$ is the Fisher information matrix. The test statistic W_k in (2) can be used to check if a change in densities has occurred at some time instant $\tau \leq k$. Under the alternative hypothesis, i.e., event at time τ, this statistic drifts for $k \geq \tau$ with the size of the drift proportional to the rate at which the test statistic moves in the direction of the alternative density. Moreover, in order to make decisions after n observations have been obtained, we use the following result (Darling, Erdos [13])

$$\lim_{n\to\infty} P\{a(\log(n)) \max_{1<k\leq n} k^{-1/2}W_k \leq t + b(\log(n))\} = \exp(-e^{-t}) \tag{3}$$

where $a(x) = (2\log(x))^{1/2}$ and $b(x) = 2\log(x) + 0.5\log(\log(x)) - 0.5\log(\pi)$. To make use of this result we set a false alarm rate f, i.e. $f = 0.001$, where $1 - f = \exp(-e^{-t})$ and we compute the threshold

$$\begin{aligned} T(f) = (2\log(\log(n)))^{-1/2}\,[&-\log(-\log(1-f)) + 2\log(\log(n)) \\ &+0.5\log(\log(\log(n))) - 0.5\log(\pi)] \end{aligned} \tag{4}$$

Then, we conclude that the alternative hypothesis is supported by the data at the first k, if

$$k^{-1/2}W_k \geq T(f) \tag{5}$$

If no such k exists for $k \leq n$ we do not reject the null hypothesis. For $n = 900$ and the two indicative values of $f = 0.01$ and $f = 0.001$ we obtain $T(f) = 4.1$ and $T(f) = 5.3$ respectively. In what follows we assume that all measurements X_i, $i \geq 1$ are independent normal random variables. In this case the test statistic in (2) is considerably simplified. Let

$$f(x_i;\mu_0,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_i-\mu_0)^2/2\sigma^2}$$

and under the alternative hypothesis (event occurence)

$$f(x_i;\mu_F,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_i-\mu_F)^2/2\sigma^2}$$

where $\mu_F > \mu_0$. The only known parameter is μ_0 that is the average value in the absence of the event.

Let $Y_i = X_i - \mu_0$ and $\mu_d = \mu_F - \mu_0$. It is clear that in the absence of event $Y_i \sim N(0, \sigma^2)$, whereas under the alternative hypothesis $Y_i \sim N(\mu_d, \sigma^2)$. In this case the test statistic is

$$k^{-1/2}W_k = k^{-1/2}\frac{\sum_{i=1}^{k} Y_i}{\left(\sum_{i=1}^{k} Y_i^2/k\right)^{1/2}} \tag{6}$$

Under the alternative, the drift of $k^{-1/2}W_k$ after a change at time τ is

$$\text{Drift of Statistic} = k^{-1/2}\frac{(k-(\tau-1))\mu_d}{(\sigma^2 + \frac{k-(\tau-1)}{k}\mu_d^2)^{1/2}} \tag{7}$$

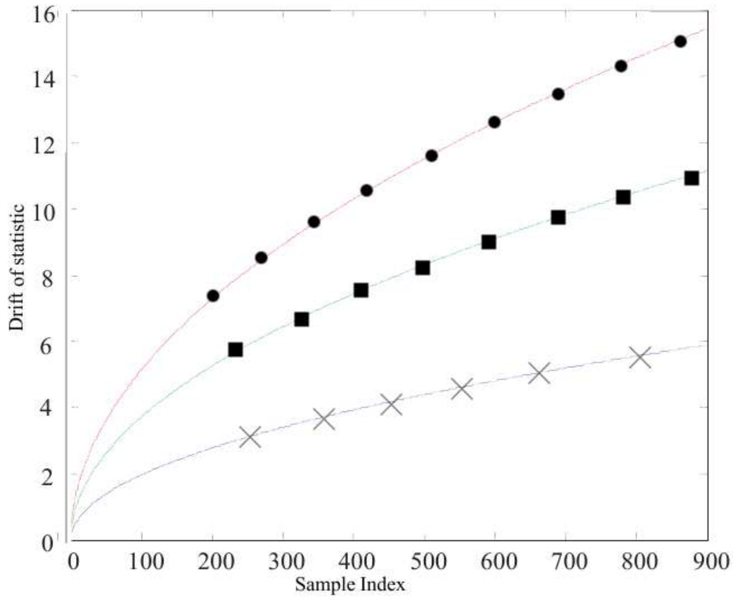

Fig. 1. Drift of statistic for various values of μ_d

Figure 1 shows the drift for $\tau = 1$, $n = 900$, $\sigma = 5$, $\mu_d = 1$ (blue-cross), $\mu_d = 2$ (green-square), and $\mu_d = 3$ (red-circle). As it is observed the greater the excess value ($\mu_d = \mu_F - \mu_0$) the largest the slope of the drift.

Table 1 shows the time instants that the test statistic crosses the thresholds of $T(f) = 4.1$ and $T(f) = 5.3$ provided that the change occurred at $\tau = 1$. As it is observed from Table 1, if the mean value excess of the alternative density is $\mu_d = 3$, it will be detected after 100 samples with a false alarm rate of 0.001.

Table 1. Threshold cross time instants

$T(f)$	$\mu_d = 1$	$\mu_d = 2$	$\mu_d = 3$
4.1	435	120	65
5.3	740	205	100

The parameter μ_0 is estimated every T_0 based on all sensors in certain area. The period T_0 should be large enough to apply the sequential detection with as many samples as possible but small enough in order to capture the frequent changes of data.

2.2 Enhancement of the Detection Mechanism

Several issues arise when the previously described detection process is adopted. First of all, the method assumes that all sensors are calibrated. It will be a problem if one of the sensors (S_i) presents a relatively large positive offset compared to the rest of the sensors. What happens in this case is that μ_0 measured at the start of the time window T_0 is constantly smaller than the measurements of sensor S_i and, therefore, this sensor will falsely indicate an event after some time, depending on the size of the offset. A remedy to this problem is the periodic calibration of the sensors. During certain periods when exogenous parameters have no effect in the sensor measurements, offsets may be calculated and taken into account in the detection process. Thus, if a sensor presents an offset of μ_{off} compared to the average value, then the detection process of this sensor will use the value $\mu_0 + \mu_{\text{off}}$ instead of μ_0.

A second issue is the correlation of the measurements. Criterion (6) was developed under the assumption that measurements are independent Gaussian distributed random variables. However, in real life measurements are correlated and this may cause a problem as shown in the following scenario. At the start of the interval T_0, when the average value μ_0 is calculated various exogenous parameters may result in an underestimation of μ_0. When such parameters do not exist anymore the average value of data will naturally increase and it will remain higher than its initial value for several samples. Depending on the relative increase and the correlation window the detection thresholds may be falsely crossed. In order to quantify and simulate the aforementioned situation we consider the following model:

We assume that sensor measurements X_i, are written as

$$X_i = \mu_0 + z_i + r_i \tag{8}$$

where z_i represents the noise due to the sensor's electronics and can be modeled as a Gaussian process of zero mean and variance σ_z^2. The random variable r_i is the sample at time i of a process $r(t)$ which models the data readings variations due to exogenous parameters. We assume that this process is Gaussian having an autocorrelation function of the form

$$R_r(\tau) = E[r(t)r(t+\tau)] = \sigma_m^2 e^{-\alpha|\tau|} \tag{9}$$

The smaller the constant α, the greater the correlation between successive samples. The process $r(t)$ can be generated by passing white Gaussian noise $w(t)$ through a system with one pole at α, that is

$$\frac{dr(t)}{dt} = -\alpha r(t) + w(t) \tag{10}$$

where the autocorrelation function of $w(t)$ is $R_w(\tau) = 2\alpha\sigma_m^2\delta(\tau)$. From equation (10) we have

$$\frac{d}{dt}\left(e^{\alpha t} r(t)\right) = e^{\alpha t} w(t) \Longrightarrow \int_t^{t+T_s} \frac{d}{dt}\left(e^{\alpha t} r(t)\right) = \int_t^{t+T_s} e^{\alpha \tau} w(\tau) d\tau$$

or else

$$e^{\alpha(t+T_s)}r(t+T_s) - e^{\alpha t}r(t) = \int_t^{t+T_s} e^{\alpha\tau}w(\tau)d\tau$$

Evaluating the previous at $t = iT_s$ we obtain

$$r_{i+1} = e^{-\alpha T_s}r_i + w_i \tag{11}$$

where

$$w_i = \int_{iT_s}^{(i+1)T_s} e^{-\alpha((i+1)Ts-\tau}w(\tau)d\tau$$

The random variable w_i is Gaussian with zero mean and variance

$$\sigma_{w_i}^2 = E[w_i^2] = \sigma_m^2(1 - e^{-2\alpha T_s})$$

Figure 2 shows a sample function of r_i which was obtained for $\alpha = 1/120$sec, $T_s = 2$sec and $\sigma_m = 1$. As can be seen from Figure 2, even small values of σ_m^2

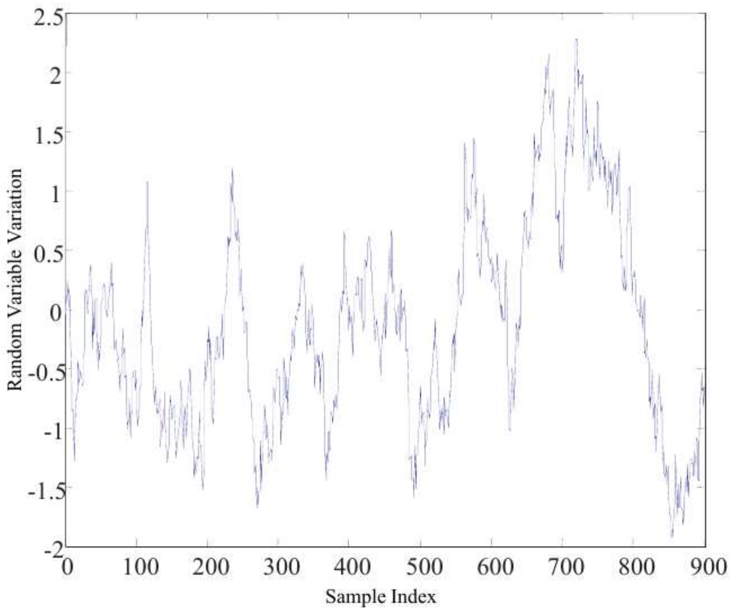

Fig. 2. Sample function of the variation random variable r_i modeling the temperature variations due to clouds, etc.

can cause large deviations from the zero mean value. The choice of the constant α indicates an average correlation window of 120 sec, that is deviations are persistent for 60 and more samples.

The problem introduced by the correlated measurements may be circumvented in one of the following ways:

1. One method is to increase the thresholds so that temporal crosses due to correlation will be avoided. A good practice is to rely on real data to set up the thresholds. However, this increase of thresholds may postpone the event detection or even cause a miss once the cross is outside the time window T_0. Increasing the thresholds will work properly only if the assumed excess mean value μ_d is quite large.
2. A more promising solution is to rely on the cooperation of neighboring sensors to minimize correlation. As the measurements of nearby sensors undergo the same variations the term r_i can be estimated from neighboring nodes and subtracted from X_i.

Based on the second approach in order to deal with the correlated measurements, consider a sensor S_i and its neighbors S_j, $j = 1, \ldots, |S_i|$, where $|S_i|$ denotes the cardinality of the neighbor set of sensor S_i. We also assume that sensor S_i senses an increase in the average value of data that is

$$X_i = \mu_0 + \mu_d + z_i + r_i \tag{12}$$

For the neighboring sensors we assume that their measurements are of the form

$$X_j = \mu_0 + z_j + r_{i-D_j} \tag{13}$$

where the noise term z_j is independent of z_i, and the term r_{i-D_j} expresses the same variation r_i that the measurements of sensor S_i undergo, delayed or advanced by D_j. Then, for sensor S_i we apply the proposed test statistic on the data.

$$Y_i = X_i - \frac{1}{|S_i|}\sum_{j=1}^{|S_i|} X_j = \mu_d + z_i + r_i - \frac{1}{|S_i|}\sum_{j=1}^{|S_i|} z_j - \frac{1}{|S_i|}\sum_{j=1}^{|S_i|} r_{i-D_j} \tag{14}$$

The term $\frac{1}{|S_i|}\sum_{j=1}^{|S_i|} z_j$ will be close to zero whereas the term $\frac{1}{|S_i|}\sum_{j=1}^{|S_i|} r_{i-D_j}$ acts as a predictor to r_i and, therefore, it almost cancels this term. Note that in applying the test statistic on data Y_i we do not have to subtract μ_0 since this term has already been cancelled.

3 Simulation results

In what follows, we will present some simulation results based on hypothetical scenarios, which emphasize the potential of the CUSUM test for the early detection of hazardous phenomena and particular fire detection. We assume two states: NOTIFY and ALERT. In the NOTIFY state the system is notifying about a possible change in the temperature mean value signaling a probable threat of a fire event. In the ALERT state, the system has to be notified on a sufficient belief for a fire event. Hence, we may use the results of Table 1 and select a false alarm rate of $f = 0.01$ to enter the NOTIFY state ($T(f) = 4.1$) and $f = 0.001$ to enter the ALERT state ($T(f) = 5.3$). When different types of sensors (e.g. temperature and humidity) exist, we can aggregate the decisions on a fire event made from the sensed contextual data (derived from temperature and humidity sensors) in order to conclude the occurence of a fire event.

We assume that the sampling rate is $F_s = 0.5$ Hz, that is one sample every 2 sec. We renew the estimation of the average temperature every 30 min (T_0) and therefore the time window to make a decision is $n = 30 \times 60 \times 0.5 = 900$ samples.

Scenario 1. In this scenario, a fire takes place 10 min ($\tau = 300$) after the estimation of the ambient temperature μ_0. This fire causes an average temperature increase from $\mu_0 = 30$ to $\mu_F = 32$ Celsius degrees ($\mu_d = 2$) at the measurements of one sensor and the standard deviation is taken $\sigma = 5$. Note that μ_F is an unknown parameter that affects the slope of the drift statistic change. Figure 3a shows a sample function of the measurements, whereas Figure 3b shows the evolution of the test statistic.

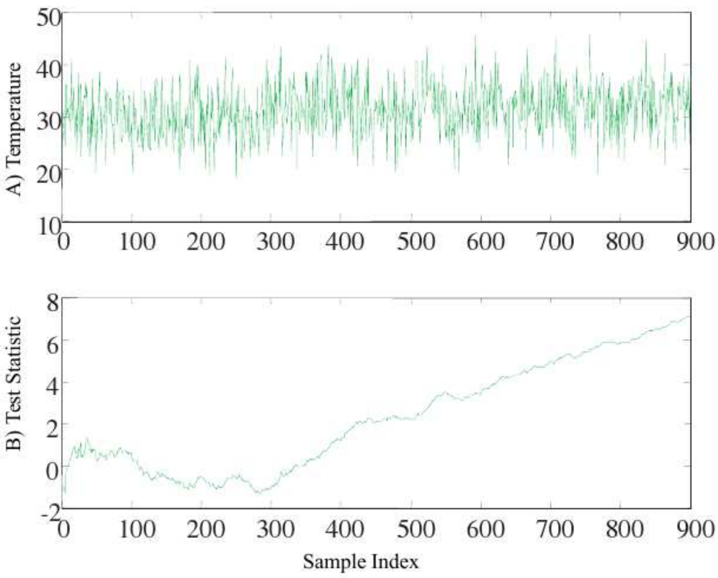

Fig. 3. (a) Sensed temperature in Celsius degree, (b) the evolution of the proposed test statistic 10 minutes after the occurence of a fire (i.e., $\tau = 300$) with an excess value $\mu_d = 2$.

Note that by using the aforementioned false alarm rates of $f = 0.01$ (threshold $T(f) = 4.1$) and $f = 0.001$ (threshold $T(f) = 5.3$), the system will enter the NOTIFY state after approximately 300 samples (10 min) and in the emergency state after approximately 400 samples (13.5 min). Note, also, that although temperature varies greatly ($20^o - 40^o$), due to the large standard deviation, the test statistic used is insensitive to instantaneous changes.

Scenario 2. This scenario simulates the case of correlated measurements. Figure 4a shows a sample function of the process X_i in Equation (12). The mean value μ_0 was set to 30^o C, $\sigma_z = 2$, $\sigma_m = 1$ and $\alpha = (120\text{sec})^{-1}$. A change of densities occurs at $\tau = 300$ with the excess mean value being $\mu_d = 2^o$C. This value is used only indicatively for simulation purposes. Higher values, make the test statistic, to drift faster and cross preset thresholds using fewer samples. Figure 4b shows the evolution of the test statistic. As it is observed from the figure the test statistic starts to drift after $\tau = 300$ but it might be that the thresholds $T(f)$ are crossed earlier, thus producing false alarms.

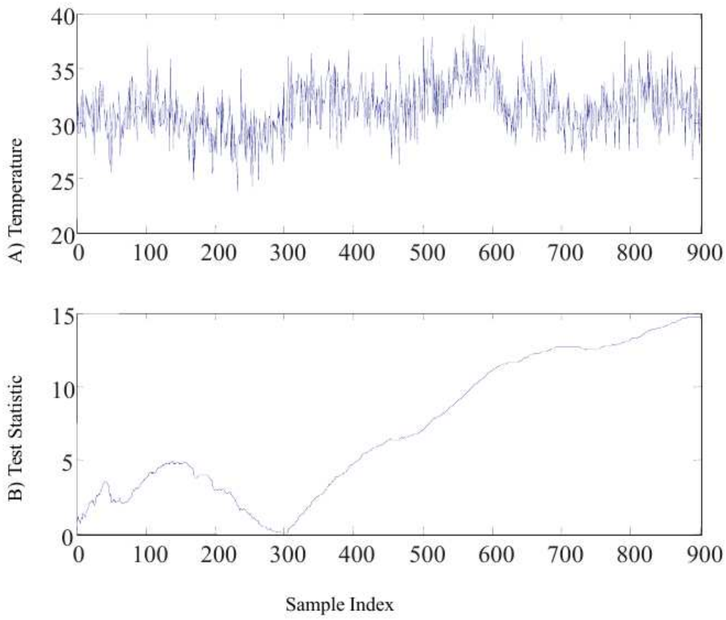

Fig. 4. (a) Sample function of the measurement process X_i, (b) the evolution of the proposed test statistic.

Figure 5 shows the simulation results for the technique proposed to mitigate correlated measurements. We assume that sensor S_i has three neighbors with corresponding delays, measured in samples, $D_1 = -3$ (time advance), $D_2 = 2$, and $D_3 = 5$. The noise z_j for each sensor is independent Gaussian with zero mean and standard deviation $\sigma_z = 2$ and $\mu_d = 2^o$ C. A change of densities occurs at $\tau = 300$ and as illustrated in Figure 5, no false crossings of the thresholds occur prior to τ.

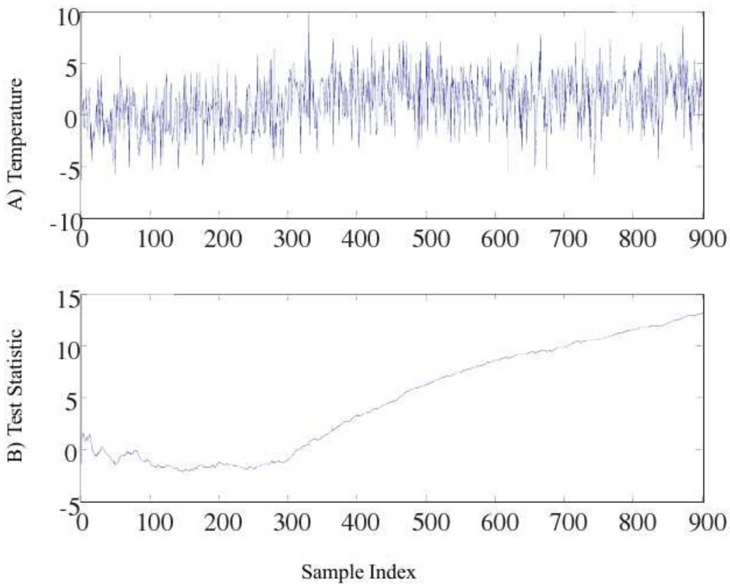

Fig. 5. (a) Sample function of the measurement process Y_i, (b) the evolution of the proposed test statistic processing the sensed contextual data from neighboring nodes.

4 Conclusions

In this dissertation we studied context management related issues for pervasive computing. A part of the dissertation deals with an event detection mechanism which is based on sensor data fusion. A cumulative sum sequential test is adopted that combines data of neighboring sensor nodes and detects changes of the underlying data distribution. The detected changes are then, compared against suitably chosen thresholds, according to a desired false alarm rate, which when crossed, the system sets it's internal notification state machine in an ALERT or a NOTIFY state. Simulation results for fire detection are also presented that verify the analysis of the proposed techniques. The synthetic trace used in our simulations contained Gaussian distributed data. CUSUM criterion is stimulated by an appropriate change of mean value of the Gaussian distribution. As a future work, we propose the enhancement of the implemented algorithms with alternative combination rules, e.g., [14],and the adoption of the Fuzzy Set theory to deal with uncertainty, imprecision and incompleteness of the underlying data.

References

1. Weiser, M.: The computer for the 21st century, Scientific American, September 1991,66-75.

2. Abowd, D.: Towards a Better Understanding of Context and Context-Awareness, Proc. International Conference of Human Factors in Computing Systems, 2000.
3. Sekkas, O., Hadjiefthymiades, S., Zervas, E,: Fusing sensor information for location estimation, in Proceedings of the 10th East-European Conference on Advances in Databases and Information Systems (ADBIS 2006), Thessaloniki, Greece, September 2006.
4. Anagnostopoulos, C., Sekkas, O., Hadjiefthymiades, S.: Context Fusion: Dealing with Sensor Reliability, in Proceedings of the 2nd International Workshop on Information Fusion and Dissemination in Wireless Sensor Networks (SensorFusion07 - MASS 2007), Piza, Italy, October 2007.
5. Sekkas, O., Anagnostopoulos, C., Hadjiefthymiades, S.: Context Fusion through Imprecise Reasoning, in Proceedings of the IEEE International Conference on Pervasive Services (ICPS 2007), pp.88-91, Istanbul, Turkey, July 2007.
6. Zervas, E., Mpimpoudis, A., Anagnostopoulos, C., Sekkas ,O., Hadjiefthymiades ,S.: Multisensor Data Fusion for Fire Detection, accepted for publication in Elsevier's Information Fusion, 2009.
7. Zervas ,E., Sekkas, O., Hadjiefthymiades, S., Anagnostopoulos C.: Fire Detection in the Urban Rural Interface through Fusion techniques, in Proceedings of the 1st International Workshop on Mobile Ad hoc and Sensor Systems for Global and Homeland Security (MASS-GHS 2007), Pisa, Italy, October 2007.
8. Sekkas O., Manatakis D., Manolakos E., Hadjiefthymiades S.: Sensor and Computing Infrastructure for Environmental Risks - The SCIER system, in Advanced ICTs for Disaster Management and Threat Detection: Collaborative and Distributed Frameworks, (Eds. Dr. E. Asimakopoulou and Dr. N. Bessis), IGI Global, November 2009.
9. Marias, G.F., Papapanagiotou, K., Tsetsos, V., Sekkas, O., Georgiadis, P.: Integrating a Trust Framework with a Distributed Certificate Validation Scheme for MANETs, in EURASIP Journal on Wireless Communications and Networking (EURASIP JWCN), Article ID 78259, 18 pages, 2006, Hindawi Pub. Corp.
10. Sekkas, O., Piguet, D., Anagnostopoulos, C., Kotsakos, D., Alyfantis, G., Kassapoglou-Faist, C., Hadjiethymiades, S., Probabilistic Information Dissemination for MANETs: the IPAC Approach, in Proceedings of the 20th Tyrrhenian International Workshop on Digital Communications, Pula, Sardinia, Italy, September, 2009
11. Page, E.S.: Continuous Inspection Schemes,Biometrika vol. 41, 1954, pp. 100-115.
12. Gombay, E., Serban, D.: An adaptation of Pages CUSUM test for change detection, Periodica Mathematica Hungarica, vol. 50, 2005, pp. 135-147.
13. Darling, D.A., Erdos, P.: A limit theorem for the maximum of normalized sums of independent random variables, Duke Math. J, vol. 23, 1956, pp. 143-155.
14. Yager,R.R.: On the Dempster-Shafer Framework and New Combination Rules, Information Sciences 41:93-137.