International Conference on Automation, Control, Engineering and Computer Science (ACECS'14)
Proceedings - Copyright IPCO-2014 ISSN 2356-5608, pp.177-180

Online autonomous monitoring in WSN with error and event detector

Aymen ABID[#1], Abdennaceur KACHOURI[#2], Adel MAHFOUDHI[#3]
[1#] *CES laboratory, ENIS, Sfax University, Tunisia*
[2#] *LETI laboratory, ENIS, Sfax University, Tunisia*
[3#] *CCIT, Taif University, Taif, Saudi Arabia*
[1]aymen.abid.mail@gmail.com
[2]abdennaceur.kachouri@enis.rnu.tn
[3]a.mahfoudhi@tu.edu.sa

ABSTRACT

Wireless sensor network (WSN) is a considerable system for sensing environmental metrics and phenomena. The condition of working for WSN e.g. harsh environment, hardware and software errors have bad affect about its sturdiness. Therefore, monitoring the network is necessary to assure the user of the data truth.
This paper presents a first step to establish a new centralized online supervisor that analyzes the dissimilarity of sensor observations to declare and confirm events and errors. Well, data are delivered to Base Station where the supervisor explores intensity of other values around a current value by vote unit and assembles spatial data in clusters. These clusters are compared with recent clusters to make decision about their state and so the current state of sensors. An experience yields awaited results using a real data base.

Keywords— Wireless Sensor Network - Autonomous Monitoring- Error & Event Detection- Data Analysis.

I. INTRODUCTION

A WSN is a platform that collects environmental observations that can be values of temperature, pressure, sound… These observations will be expedited to Base Station (BS) where the user can explore them. In this point, we will try to ameliorate the confidence of user in this current information. In fact, we will not overcharge the sensor by distributed detection but we will use an autonomous monitoring procedure to declare events and errors by detecting outlier clusters and data.

In our recent works we treat data to detect errors by voting principle [1] and auto configured computing for limits of right values and so for declaring values outside this right interval as faults using Bayesian network [2]. In this work, we treat both errors and event using grouping, clustering and vote technics. In fact, recent researches try to resolve errors and event simultaneously with an outlier data detection module [3].

In this paper we try to create clusters after a voting process in order to define both data anomalies (errors) and events with spatial outlier detector [4]. Effectively, cluttering is an outlier technique that is used to define events and anomalies in WSN [5]. It's also a technique for data mining area[5], image processing [7] and for many other applications [8].

II. MONITORING PROCEDURE

The monitoring is a process that decides the state of a sensor; it participates in event or it fails. This subsystem contains these following tasks (Fig.1):

- Data acquisition: In this activity we define the manner with the data are collected from sensor and so when the monitoring system will be executed (each period, continuous, when new alteration of sensor value…). In fact, the system tripped when a new value different from its previous is arrived.
- Initialization: It needs at least one execution with learning data to define the initial large of valid values in a cluster. The large of one accepted values set is called confidence interval (CI) [4].
- Voting: For each current sensed value, we count the values in a set of thereabout. The limits of adjacent values are defined by "CI". After that, the algorithm groups the set that has at least one common value.
- Clustering: It creates clusters from groups that
  - Not attributed to any current cluster
  - Not included in recent cluster (of last execution)
  - Doesn't include a recent cluster.
- Diagnostic: It's the activity to define sensor states by making decision for each current cluster if it's
  - Erroneous cluster: The values included in this cluster are erroneous values (sensors).
  - Correct cluster: that can be a recent event or a new event and so the end of recent cluster.

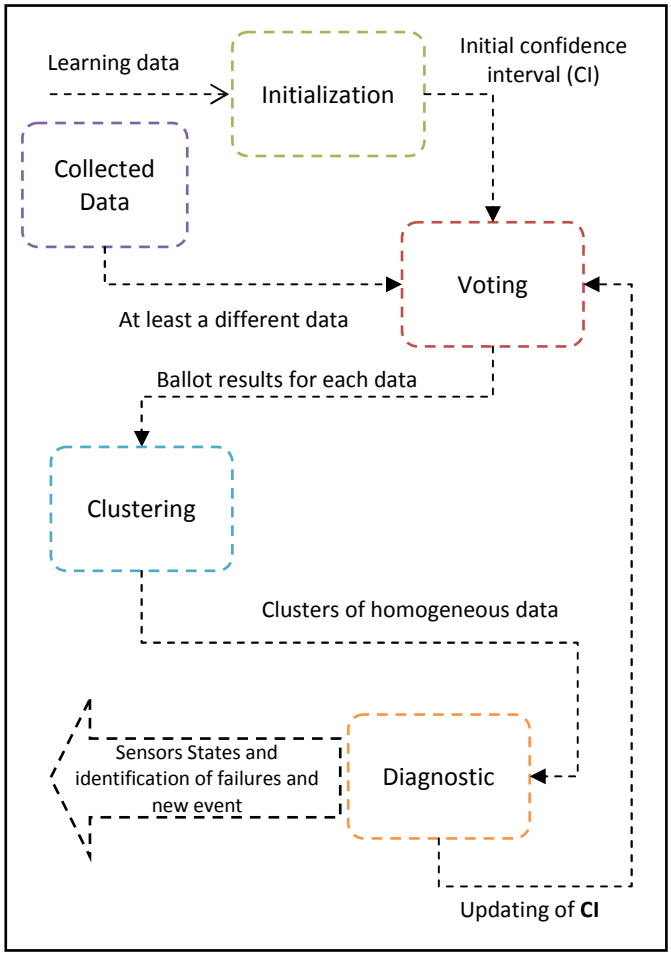

Fig.1Context diagram for monitoring system in WSN

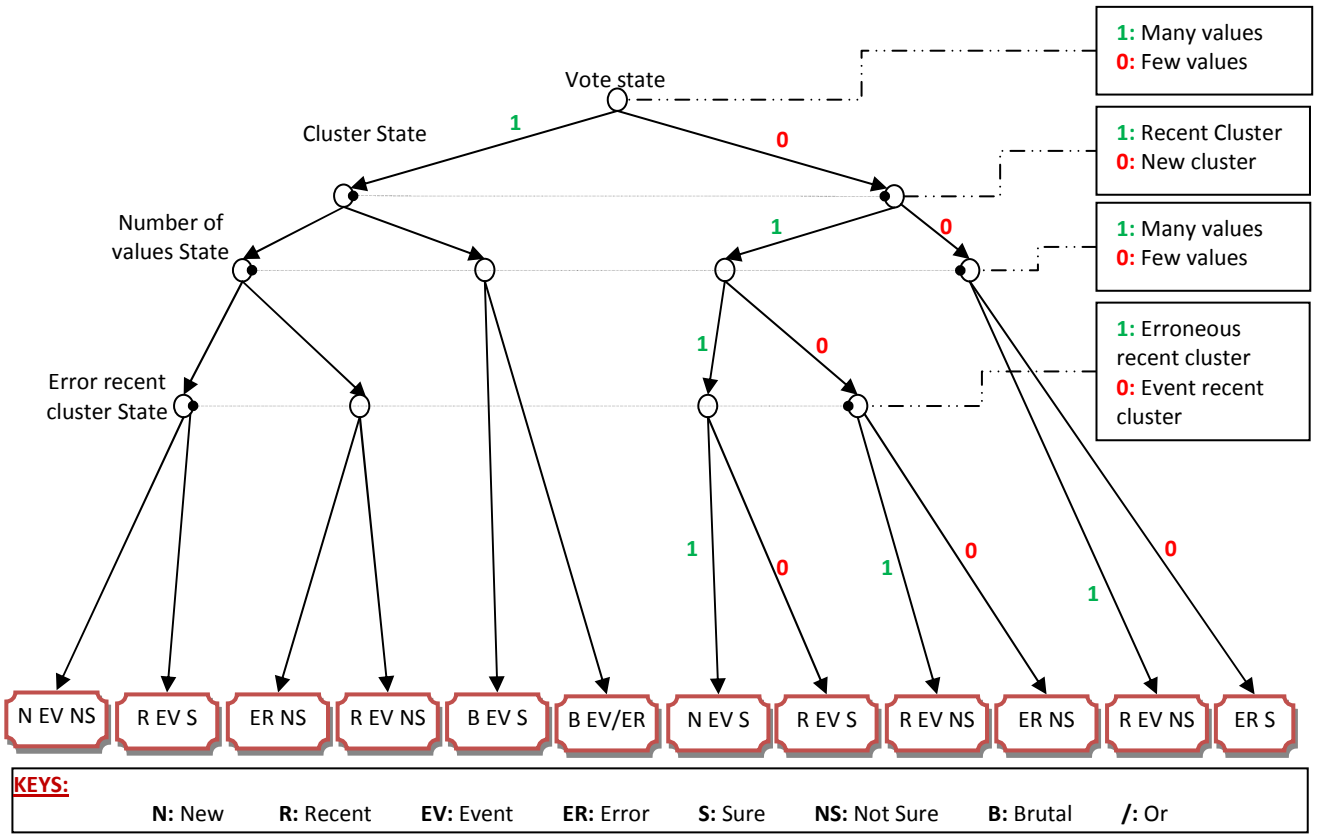

Fig.2 Binary tree of decision to decide a current cluster state

III. MAKING DECISION

Making decision will be done through diagnostic phase (fig.1). According to vote results and clustering activity the decision will be defined using a binary tree (Fig.2). In fact, this tree will give the state of current clusters (clusters that we construct from data of this current execution). The decision is to decide if a sensor is erroneous or normal. It's also to decide if there are new events or not.

IV. RESULTS

For experiment we used a temperature data base collected by 6 Telosb sensors with TinyOS system working thought Windows system using cigwin terminal [9] in 25m^2 room.

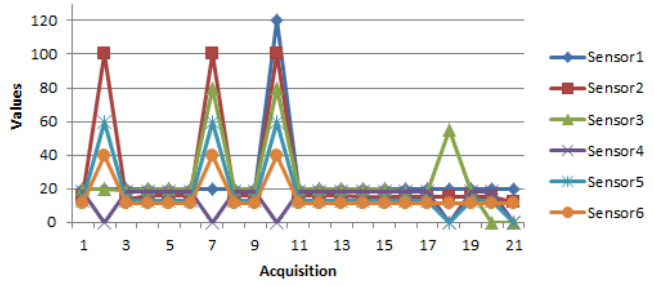

Fig.3 Temperature values from wireless Telsob sensors with the injection of some errors and events

In collected data, we introduce some heterogeneous errors and synthetic event as a second event with the real event, to have more confidence in the reaction of monitoring system. These insertions give up more extreme error or event cases to be detected (Fig.3) (Table I).

TABLE I DATA BASE USED FOR EXPERIMENT

Acquisition	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6
1	20	15	20	18	13	11
2	20	100	20	0	60	40
3	20	14	19	18	13	11
4	20	16	20	18	13	11
5	20	16	20	18	13	11
6	20	16	20	18	13	11
7	20	100	80	0	60	40
8	20	16	20	18	13	11
9	20	16	20	18	13	11
10	120	100	80	0	60	40
11	20	16	20	18	13	11
12	20	16	20	18	13	11
13	20	16	20	18	13	11
14	20	15	20	18	13	11
15	20	15	20	18	13	11
16	20	15	18	18	13	11
17	20	15	18	18	13	11
18	20	15	55	0	0	11
19	20	15	18	18	13	11
20	20	15	0	18	13	11

21	20	12	0	0	0	11

Calculating was by the following formulas:

FAR=NBFDR/NBV (1)

FPR=NBPDR/NBV (2)

With **NBFDR** is number of false detection, **NBV** is the number of values and **NBPDR** is the number of positive detection.

Results are interesting, the false alarm rate "FAR" (false declaration of errors) and false positive rate (false declaration of event) are not with huge average (Table II).

We notice that all errors are detected with this monitoring process, e.g. errors of execution, just after the first acquisition (learning phase).

After the initialization phase with the first acquisition, we introduced four erroneous values nonhomogeneous (uncorrelated). The treatment gave the results shown in the fig.4.

```
*************          INIT (Acquisition 1)  *************
S[1]=20.000000  S[2]=15.000000  S[3]=20.000000  S[4]=18.000000
S[5]=13.000000  S[6]=11.000000
Half Confidence Interval (CI)= 4.500000
*************          Acquisition 2  *************
S[1]=20.000000   S[2]=100.000000     S[3]=20.000000     S[4]=0.000000
S[5]=60.000000   S[6]=40.000000
Clusters: [15.500000;24.500000] [95.500000;104.500000]
[-4.500000;4.500000] [55.500000;64.500000] [35.500000;44.500000]
Diagnostic
Half Confidence Interval (CI)= 4.500000
cluster[1]: vot_sta=1 clu_stat=1 nb_val_sta=1 =>recent event (sure)
cluster[2]: vot_sta=0 clu_stat=0 nb_val_sta=0 =>  error (sure)
cluster[3]: vot_sta=0 clu_stat=0 nb_val_sta=0 =>  error (sure)
cluster[4]: vot_sta=0 clu_stat=0 nb_val_sta=0 =>  error (sure)
cluster[5]: vot_sta=0 clu_stat=0 nb_val_sta=0 =>  error (sure)
FINAL SENSORS STATES: 1=>Erronous 0=>Event
s[1]=0        s[2]=1        s[3]=0        s[4]=1        s[5]=1        s[6]=1
```

Fig.4 Example of results during detection processing

The decision for clusters implies that all values that belong to this cluster have had this decision. So the sensors 2, 4 and 6 are reported incorrect.

In reality, for the acquisition '7', only one value is correct. '0' is an incorrect value which will in another time a trigger value of an event. This choice is in order to put in the worst case.

For execution '10', all sensors fail to give correct value. In this case, the monitoring system declared all values as not sure events and it was not the case.

As possible remediation, we can introduce the ancient management of values and the preceding statements. In this context, that is made in the acquisition '17 'and '18' will give the needed help.

For execution '18', the false decision is due to the apparition of new event with one value at first; it is sensed with only one sensor.

The acquisition '20' has an event but by the presence of a single value. As a result, it is difficult to discover it. But this was recovered in the next run.

In this test, "RPF" and "NSDR" have the same grade because the system was unable to give a justified assessment when the network has produced a major disruption. In fact, almost all the values are wrong. In other cases, it may be different, e.g. set of errors correlated can be viewed as event. 'FAR' could be lower (in the order of 0.008) if acquisition '20 'is actually stated as the beginning of an event.

TABLE II ACCURACY RESULTS

Acquisition	FAR	FPR
2..6	0	0
7	0	0.833
8..9	0	0
10	0	1
11..17	0	0
18	0.166	0
19	0	0
20	0.166	0
21	0	0
Average	0,017	0,092

V. DISCUSSION AND CONCLUSIONS

In this paper, we develop an online process to detect errors and events. This monitoring procedure presents good results for FAR and FPR. But, the confidence in the making of decision can be more important when we ameliorate the diagnostic phase.

What is expected in the next work is to add other criteria for decision-making as energy sensors status, link status, location of sensors, history of values, logs of decisions of the detector sensor and other information... Then, we can develop an algorithm to initialize parameters with learning phase. In the evaluation level, it will be appropriate to compare this detection procedure with other work using the same data base and the same network configuration (topology, communication, routing ...). It is preferable to use a range of values containing real errors and events. These anomalies cloud be the result of several phenomena as fire, intrusion, failure, congestion... A synthetic base may also justify the robustness of the detector. Indeed, injecting events or incorrect values for an actual base may complicate the situation and gives more difficulty to properly data analysis. However, a good formulation of an error or event is another level of difficulty to consider.

VI. REFERENCES

[1] A. Abid, H. Kaaniche, A. Kachouri, M. Abid, *Quality of service in Wireless Sensor Networks through a failure-detector with Voting Mechanism*, ICCAT, IEEE, Jan.2013.

[2] A. Abid, M. A. Maalej, A. Kachouri, A. Mahfoudhi, "*A Bayesian Network for an auto configuration limits of valid values in WSN failure detection*", ROIS, Sep2013.

[3] M. Zakirul, et Al., "Localized Decision-Making in Wireless Sensor Networks for Online Structural Health Monitoring", Central South University School of Information Science and Engineering, Technical Report, No. TR-SISE-02:1–55, 2013.

[4] Zhang, Y., Hamm, N. A., Meratnia, N., Stein, A., van de Voort, M., & Havinga, P. J. (2012). Statistics-based outlier detection for wireless sensor networks. International Journal of Geographical Information Science, 26(8), 1373-1392.

[5] Lim, T. H. (2010). Detecting anomalies in Wireless Sensor Networks. Qualifying Dissertation, University of York.

[6] Berkhin, P. (2006). A survey of clustering data mining techniques. In Grouping multidimensional data (pp. 25-71). Springer Berlin Heidelberg.

[7] Celenk, M. (1990). A color clustering technique for image segmentation. Computer Vision, Graphics, and Image Processing, 52(2), 145-170.
[8] Sneath, P. H., & Sokal, R. R. (1973). Numerical taxonomy. The principles and practice of numerical classification.
[9] http://pequan.lip6.fr/~bereziat/cygwin/