International Journal of Computer Networks and Communications Security

VOL. 5, NO. 5, MAY 2017, 105–114
Available online at: www.ijcncs.org
E-ISSN 2308-9830 (Online) / ISSN 2410-0595 (Print)

Adaptive Scheme for Outliers Detection in Wireless Sensor Networks

Blaise Omer Yenké[1], Moussa Aboubakar[1], Chafiq Titouna[2, 3], Ado Adamou Abba Ari[2, 4] and Abdelhak Mourad Gueroui[2]

[1] LASE Laboratory, University of Ngaoundéré, CAMEROON

[2] LI-PaRAD Laboratory, Université Paris Saclay, University of Versailles Saintt-Quentin-en-Yvelines, FRANCE

[3] Department of Computer Science, Faculty of Mathematics and Computer Science, University of Batna 2, ALGERIA

[4] Department of Mathematics and Computer Science, Faculty of Science, University of Maroua, CAMEROON

ABSTRACT

Nowadays, Wireless Sensor Networks (WSNs) have become a way of information gathering and monitoring. These particular kinds of ad hoc networks find applications in various fields such as healthcare, army and environment, etc. However, WSNs are subjected to a number of faults, either due to the worst communication links, the hostile environment in which they are deployed or due to energy constraints. Thereby, to ensure the Quality of Service (QoS) in such networks, it is important that WSNs become capable of detecting and recovering erroneous data. In this way, many outlier detection techniques have been proposed. These techniques are based on information theory, statistics and other techniques. The Nearest Neighbour Search is one of these techniques. It uses the Euclidean or Mahalanobis distance for detecting outliers in a given sensor networks. In this paper, we proposed a distributed adaptive approach for the detection of outliers. The proposed approach is based on Euclidean or Mahalanobis distance, depending on the size of the data collected. Extensive experiments have been conducted and the results confirmed the effectiveness of the proposal.

Keywords: *Wireless Sensor Networks, Outlier Detection, Mahalanobis Distance, Eucidean Distance, Fault Tolerance.*

1 INTRODUCTION

A. Background

A Wireless Sensor Network (WSN) can be considered as a network of small wireless sensor devices denoted as nodes, which can sense information from a given environment and communicate the gathered data through wireless link to a sink, which can locally exploit the data or forward it to end users through gateways [3,28,31]. Moreover, WSN have infiltrated our daily life such as medical monitoring, military surveillance, vehicle monitoring, home automation monitoring, weather monitoring, building structures monitoring, and industrial plant monitoring [12, 16, 25].

However, in some applications, nodes are deployed in remote and harsh environments (forest fire, earth- quake or chemical spill) [4,24]. In such areas, nodes can be failed due to the energy depletion, hardware failures, communication link errors and even intrusion from attackers [13,23]. These problems reduce the quality of the gathered data and the entire network. At this stage, it is necessary to set up a mechanism to ensure the quality of the collected data in order to allow taking suitable decision. To improve the quality of data in WSN, we should remove all the abnormal data from the network. This incorrect one is called outlier, which is defined as an observation that deviates from other observations as to arouse suspicion that it was generated by a different mechanism [9]. Many techniques are proposed in literature for detecting outliers in a WSN. They are categorized in [30]:

– Statistical-based approach;
– Nearest Neighbor approach;
– Clustering based approach;
– Classification-based approach;
– Spectral Decomposition based approach.

We also note hybrid approaches which combine at least two techniques. This allows in some cases to exploit advantages of approaches while overcoming their weaknesses. Most of the existing works did not take in account multivariate data and assumed that the data are univariate [23,30]. In fact, they did not consider when the attributes are together. This fact can create anomaly while we do not observe any problem when we have an individual attribute.

B. Authors' contributions

In this paper, we present an adaptive model for outlier detection in WSN that is based on Nearest Neighbor approach. We use Mahalanobis and Euclidean distance inside clusters for detecting outlier in a set of gathered data. The main idea behind our technique is to exploit one of these two theorems, depending on the size of the data collected by sensor nodes, for detection of outliers. In short, our main contributions can be summarized as follows:

– formulation of the outlier detection problem;
– determining value of transition between two metrics for outliers detection;
– proposition of time for transition between two detection;
– proposition and evaluation of two metrics for outliers detection in WSN;

C. Organization of the paper

The rest of the paper is organized as follows. Sect. 2 presents a brief review of some related works with the outliers detection in WSNs. Sect. 3 presents the system design of our proposal and the problem formulation. Sect. 4 describes the proposed approach and the scheme design while in Sect. 5 we show the performance evaluation and a discussion on results. Finally, Sect. 6 reports conclusion of our work and the future direction.

2 RELATED WORKS

During the past few years several techniques for detecting outliers in a WSN have emerged to improve the quality of the data collected by the sensors and thus allow their proper use. These techniques are those based on: statistical, nearest neighbor, spectral decomposition, clustering, classification [10, 30, 15, 2].

In statistical based approach, the authors in [29] addressed the quality of data in WSN by proposing a technique which can identify outlier in a real time manner. The methodology proposed makes use of time-series analysis and geostatistics. *Bettencourt et al.* [5] proposed a distributed method for anomalies identification and detection in ecology. In their approach, each sensor node examines the statistical distribution between its own measures and each of its neighboring nodes, and between its past and existing measures.

However, the techniques described above have in common that the aim is to detect the error and to distinguish between error and event. A drawback of these techniques is that they do not take into account the energy aspect which is also the cause [20] of outliers. *Sheng et al.* [21] proposed a Histogram based technique to identify anomalies. The authors claim that their technique can reduce significantly the communication overhead.

Moreover, in Nearest Neighbor based approach, *Salem et al.* [19] proposed a technique for detecting abnormalities in health field, particularly in the context of monitoring vital signs of patient. The proposed method use Mahalanobis distance for the detection of outliers and Kernel density to identify the source of the outlier. This approach aims at making the difference between a wrong measurement and effective degradation of patient's condition. The authors in [11] proposed a technique for identifying a subset of outlier among all data collected by the sensors. The Euclidean distance and fuzzy logic is used for the detection of outlier. With this technique outlier is firstly identified and secondly classified as event or error.

However, the nearest neighbor approach has a high computational complexity as mentioned in [10, 20]. This is an inconvenience for techniques based on this approach. Some research work has been done to alleviate this problem. *Xie et al.* [27] proposed a hypergrid *K*-NN (Nearest Neighbor) online scheme for outlier detection. This scheme aims at reducing computational complexity and communication overhead.

In spectral decomposition, *Mahmoud et al.* [14] proposed a method to find abnormal behavior of user activity in smart home environment equipped with sensors. The method uses Principal Component Analysis (PCA) and fuzzy-rule based for identification of abnormal behavior in an intelligent environment and then classifies them. In [8], authors proposed a centralize scheme for outlier detection. The scheme makes use of

Principal Component Analysis to resolve the problem of data integrity and accuracy. Most techniques for outlier detection which are based on spectral decomposition use PCA. However according to [30], selection of suitable principle components is computational expensive.

In clustering based approach, *Aggarwal and Singh*

[1] proposed a technique that uses *K*-mean algorithm for clustering and then find outlier dataset. Manhattan and Euclidean distance are used in that algorithm to find mean of each cluster. The mean is then used to detect outlier. Furthermore, *Rajasegarar et al.* [17] proposed a distributed anomalies detection which aims at minimizing communication overhead. Moreover, *Branch et al.* [7] have also proposed an approach for the detection of outliers in a WSN. This approach allows both a reduction in the use of the bandwidth and a reduction in energy consumption. It uses a single jump for communication, which facilitates the detection of faulty sensor nodes. However, clustering based technique suffers from the choice of an appropriate parameter of cluster width [30].

Moreover, classification based approaches are divided into Support Vector Machine (SVM) based and Bayesian Network based. In Bayesian network based, *Chafiq et al.* [22] proposed an approach that uses Bayesian and inference concept to well classify outliers and normal data. This approach improves the accuracy of detection by the use of two levels (the sensor nodes and the cluster head or gateway). The idea behind this technique is to exploit the hierarchical structure of the network and the limited computing capacity of sensor nodes. At the first level, the sensors can tell whether or not they detect erroneous data and the decision of each sensor node will be merged on a large scale (at the gateway, cluster head) to detect erroneous data. This merger provides support for dynamic nature of networks.

In addition, *Salem et al.* [18] proposed a framework to identify abnormal behavior in wireless body area network for healthcare monitoring. The proposed method uses SVM to classify abnormal data in incoming sensor data. This approach has been conducted using real patient data and the results show that anomalies are quickly identified but the communication overheads are not taken into account.

In this paper, we propose an adaptive outlier detection model for wireless sensors networks. The model uses Mahalanobis and Euclidean distance for detection while taking into account the size of data gathered by the sensors. The model is performed in a distributed manner and aims at taking into account the correlation between the gathered data.

3 SYSTEM DESIGN

A. Assumptions

We consider a model of WSN with N sensors scattered in a *2*-dimensional environment. Each sensor has a unique identifier $i \in [1, N] \cap$. We assume that all sensors are stationary and homogeneous. In addition, their batteries cannot be recharged. Moreover, sensors have the same computation and power capabilities. We assume that our network is organized in clusters and let us consider that all the gathered data are collected by Cluster Heads (CH) and forwarded to a central Base Station (BS), where the data processing occurs. However, we assume that some local processing may occur in sensor nodes as well as in CHs in order to reduce the overall communication cost. Finally, let us consider that there are no malicious attacks on the considered WSN.

B. Problem formulation

The proposed outliers detection technique is a distributed adaptive model based on the Nearest Neighbor Search (NNS) [6]. The NNS is seen as a proximity search problem that is brought back to an optimization problem for finding closest points on a *d*-dimensional space. In fact, the main goal of the NNS is to find a point that minimizes a given objective function, which is in our case, the distance between vectors of gathered data by sensors. Formally, let M be a *k*-dimensional subspace of ^{d} ,i.e., $M \subset {}^{d}$. Let $V = \{v_1, v_2, .., v_n\}$ be a set of n data vectors all living in M and having at least unit norm. Also, let us consider a query vector $q \in M$. The resulting optimization can be formulated as given in Equation 1.

$$X_{\rho} = \underset{v_i \in V}{argmin} \rho(v_i, q) \qquad (1)$$

Generally, M is a metric space and the considered dissimilarity is a distance metric. Hence, by considering the Equation 1, a simple solution to the NNS problem is to compute ρ as a distance measure from the query data vector q to each sensor's data vector in V and return the closest one. This approach is known as the linear search method and usually it is guaranteed to find the exact nearest data vector. Moreover, since M is taken to be the d-dimensional vector space, in our proposal he dissimilarity is measured by using the Euclidean distance and the Mahalanobis distance in order to detect outliers while taking into account the size of the gathered information by each sensor, with a maximal computational complexity of $O(n \times d)$.

4 PROPOSED APPROACH

A. *Detection metrics*

At each sensor level, we used either the Euclidean distance or the Mahalanobis distance metrics to detect outliers. Formally, the Euclidean distance $d_E(v_i, v_j)$ between two data vectors $v_j = (v_{i1}, v_{i2}, ..., v_{id})^T$ and $v_j = (v_{j1}, v_{j2}, ..., v_{jd})^T$ $i, j \in [1, N] \cap$, in the *d*–dimensional space ^{d} is defined in Equation 2.

$$d_E(v_i, v_j) = \sqrt{(v_i - v_j)^T (v_i - v_j)} = \sqrt{\sum_{k=1}^{d} \left(v_{i_k} - v_{j_k}\right)^2} \tag{2}$$

Finally, for a given threshold of the data, the correlation between variables is needed in the computation of the distances among data vectors. The correlation need to be considered since there are associations among data vectors. Used for the detection of outliers, the Mahalanobis distance $d_M(\vec{v}_i, \vec{v}_j)$ can be defined as a dissimilarity measure between two random data vectors $v_j = (v_{i1}, v_{i2}, ..., v_{id})^T$ and $\vec{v}_j = (v_{j1}, v_{j2}, ..., v_{jd})^T$, $i, j \in [1, n] \cap$, in the *d*-dimensional space ^{d} of the same distribution with d X d covariance matrix Σ. The Mahalanobis distance $d_M(\vec{v}_i, \vec{v}_j)$ is given in Equation 3.

$$d_M(\vec{v}_i, \vec{v}_j) = \sqrt{(\vec{v}_i - \vec{v}_j)^T \Sigma^{-1} (\vec{v}_i - \vec{v}_j)} \tag{3}$$

Where Σ^{-1} is the inverse of the covariance matrix Σ.

In particular, if a data vector $\vec{v}_j = (v_{j1}, v_{j2}, ..., v_{jd})^T$ has a vector $\mu = (\mu_1, \mu_2, ..., \mu_d)^T$ of means of the *d* data vectors and if the underlying distribution of the *d* random data vectors is exactly multivariate normal with a $d \times d$ covariance matrix Σ, then Mahalanobis distance $d_M(\vec{v}_i)$ of a given data vector $\vec{v}_i$ from μ can be defined as given in Equation 4.

$$d_M(\vec{v}_i) = \sqrt{(\vec{v}_i - \vec{u})^T \Sigma^{-1} (\vec{v}_i - \vec{u})} \tag{4}$$

B. *Scheme design and algorithm*

In the design of the proposed scheme, we take into account the correlation between the small size data. In addition, our proposal aims at reducing the computation complexity introduced by large size data. Concretely, within clusters, each sensor node will itself decide on detecting an outlier by using our proposed NNS scheme. The detection of outlier is achieved by an alarm decision function. Let us define a value γ that represents the exact number of gathered data by a sensor in a given data vector. So, by supposing that $\chi \sim N_d(\mu, \Sigma)$, $\chi = (\chi_1, \chi_2, ..., \chi_d)$ is a random d -dimensional random data vector that follows a multivariate normal distribution with the mean vector μ and a positive $d \times d$ covariance matrix Σ, then the $d_M^2 \sim \chi_{\gamma, 0.975}$, where 0.975 is the quantile used for the detection of outliers. So, the alarm decision function is given in Equation 5 and the associated detection algorithm is described in Algorithm 1.

$$\mathcal{A}_i = \begin{cases} 1 & \text{if } d_M, d_E \geq \sqrt{\chi^2_{\gamma, 0.975}} \\ 0 & \text{Otherwise} \end{cases} \tag{5}$$

Algorithm 1 . The alarm signaling algorithm

Function *alarmSignaling*

Begin

1. *Input*:
Data vectors
1.1. $u = (u_1, u_2, ..., u_d)^T$
1.2. $v = (v_1, v_2, ..., v_d)^T$
Distance Metric
1.3. _MAHALANOBIS_ or _ECLUDIAN_

2. *Outlier Detection*:
2.1. Compute the distance according to Equations 2 or 3.
2.2. Compute the alarm signaling $\mathcal{A}_i$ according to Equation 5.
2.3. **if** $\mathcal{A}_i == 1$
2.4. *Return* _OUTLIER_
2.5. **else**
2.6. *Return* _INLIER_

End

In the proposed scheme for outliers detection, the data gathered by the sensors are initialized in a d-dimensional data vector $v = (v_1, v_2, ..., v_d)^T$. Let θ be a certain threshold used to choose the accurate distance measure ρ that is used for determining outliers, as described in the following points. From the large experiments conducted with γ less than 100, we got θ = 50. Hence:

- if $\gamma < \theta$, then we compute the Mahalanobis distance between the recordings of a sensor, i.e., v and that of one of its neighbors. If the square of the distance is greater than $\chi^2_{\gamma, 0.95}$, then we are in the presence of an outlier otherwise it is an inlier, i.e., a normal value.
- if $\gamma \geq \theta$, then we compute the Euclidean distance between the recordings of a sensor and that of one of its neighbors. If the square of the distance is greater $\chi^2_{\gamma, 0.95}$, then we are in the presence of an outlier otherwise it is an inlier.

Furthermore, our outliers detection scheme

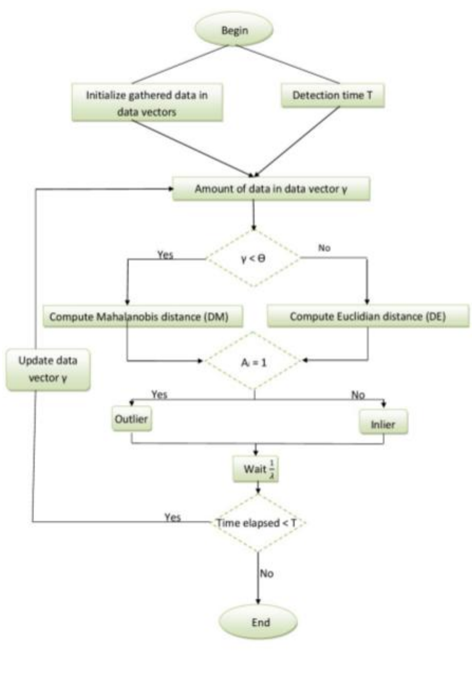


Fig. 1 Flow chart of the proposed scheme

follows a Poisson process with parameter λ. So, we used the exponential distribution to determine the time between successive outliers detection. By supposing that x is the time between detection, the probability density function is given in Equation 6. Without loss of generality, we consider a pause time $\bar{t}$ (see Equation 7) that the mean time between detection.

$$f(x) = \lambda e^{-\lambda x} \tag{6}$$

$$\bar{t} = \frac{1}{\lambda} \tag{7}$$

The complete steps of the proposed scheme is given in the Algorithm 2. Also, in order to have a clear picture on operations achieved in the proposal, Figure 1 provides a flow chart showing the whole operations performed in the proposed scheme.

Algorithm 2 . The outliers detection algorithm

```
Begin

Initialize data vector;
Initialize T;
Set TimeElaspsed = 0;
do
    for each cluster
        for each sensor
            if γ < θ
                Launch alarmSignaling with
                _MAHALANOBIS_
                Wait a time t̄ before the next detection.
            else if γ ≥ θ
                Launch alarmSignaling with
                _ECLUDIAN_
                Wait a time t̄ before the next detection
    Update Data vector;
    Update TimeElaspsed;
while TimeElaspsed < T

End
```

5 PERFORMANCE EVALUATION

In this section, we propose a performance evaluation of the proposed scheme for outliers detection by performing extensive simulation according to some proposed evaluation metrics.

A. Evaluation Metrics

To evaluate the performance of the proposed algorithm, we analyzed the following metrics:

1) Detection Accuracy Rate

It is a metric that indicates the rate of the detection accuracy (see Equation 8). In the ideal scenario, the ratio should be equal to 1. If the ratio is very low compared to the ideal, this indicates some faults in the approach design.

$$DAR = \frac{OutlierDtected}{AllOutlier} \tag{8}$$

Where *OutlierDetected* represents all the gathered data considered as outlier and *AllOutlier* represents the total amount of outliers.

2) False Alarm Rate

This metric (see Equation 9) represents the number of false alarms, i.e., the data assumed to be false among all the correct data. To achieve a reliable detection, we should have a very small rate.

$$FAR = \frac{DataAssumedFalse}{AllCorrectData} \quad (9)$$

Where *DataAssumedFalse* represents the amount of false gathered data and *AllCorrectData* represents the amount of correct gathered data.

B. Simulation parameters

Experiments have been performed by using the realistic node behaviors, radio and wireless channel models included in the Castalia 3.2 simulator, which is based on the OMNet++ 4.6 platform. Simulations were carried out with randomly and uniformly deployed stationary sensors on a sensing area whose dimensions are $100 \times 100\ m^2$. To provide realistic results, we consider five scenarios of network with sizes of 10, 50, 60, 80 and 100. All sensors are homogeneous, there is no special sensor with a particular behavior. The key simulation parameters are depicted in the Table 1.

Table 1 Simulation parameters

Parameter	Value
Square (m^2)	100×100
Number of sensors	10, 50, 60, 80, 100
Number of clusters	1, 2, 3, 4, 5
Data packet size (Bytes)	4
Bandwidth (KByte/s)	16
Queue size (packets)	5
MAC protocol	T-MAC
Transmission channel	Wireless link
Radio layer	CC2420 Radio layer
Propagation model log	Normal path loss model
Simulation times (s)	60, 300, 600

C. Results and discussion

In the following, we present the results obtained when applying our distributed adaptive model for outlier detection. We have applied the model in WSN with and without clustering. We executed our algorithm on different periods (60s, 300s, 600s) with various number of sensors (10, 50, 60, 80, 100). To assess the performance of our model, we injected outliers at various time of the simulation in order to compute the detection accuracy rate (DAR) and the false alarm rate (FAR). The basic idea is to increase the number of outlier and observe the behavior of our model.

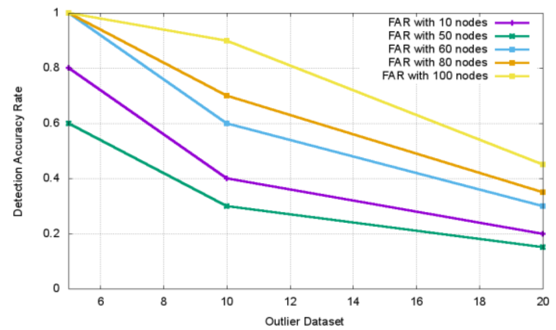


Fig. 2. Detection accuracy rate after 60 s without clustering

Figure 2 represents the detection accuracy rate obtained after 60s of simulation of our model without clustering the sensor nodes. The Detection Accuracy (DA) is reaches 1 with 60, 80 and 100 sensor nodes for 5 injected outliers dataset. The best scenario is thus reached.

In Figure 3, we report the detection accuracy obtained after simulation for our model over a period of 60s while clustering the sensor nodes. The DA reaches value 1 with 50, 60, 80, 100 sensor nodes for 5 outliers dataset. This best scenario is also achieved with 100 sensor nodes for 10 injected outliers dataset. Figure 4 presents the False alarm rate of our model.

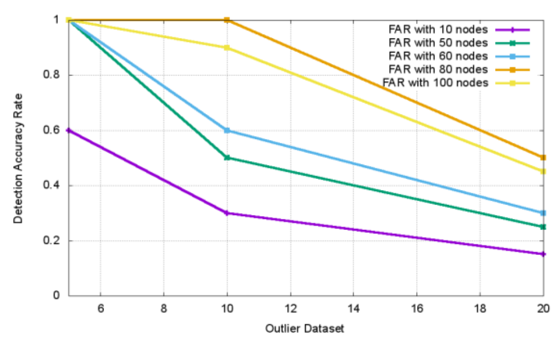


Fig. 3. Detection accuracy rate after 60 s with clustering

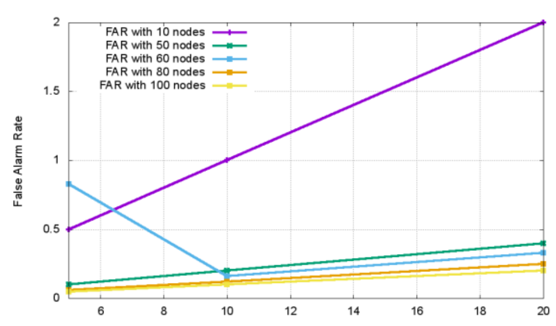


Fig. 4. False alarm rate 60 s

We repeated the simulation of our model this time over a period of 300s without clustering the sensor nodes. Figure 5 represents the results obtained. The DA reaches value 1 with 60, 80, 100 sensor nodes for 5 dataset of outliers injected in WSN. For 10 outliers dataset inject, value 1 is reached with 100 sensor nodes.

Figure 6 represents the results obtained after 300s of simulation while clustering sensor nodes. The DA reaches value 1 with 60, 80, 100 sensor nodes for 5 outliers dataset injected in WSN. It also reached 1 with 80 sensor nodes for 10 dataset of outliers injected.

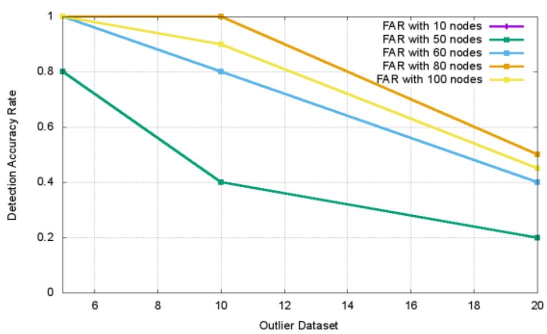


Fig. 5. Detection accuracy rate after 300 s with clustering

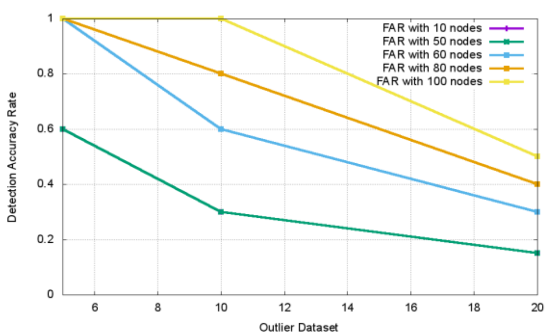


Fig. 6. Detection accuracy rate after 300 s without clustering

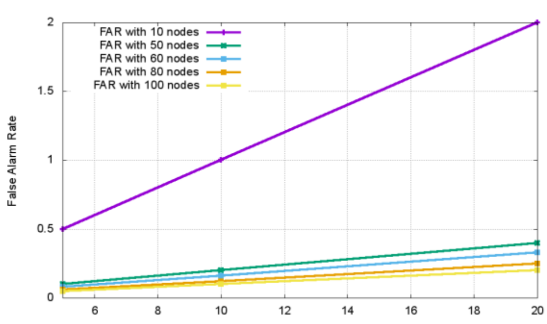
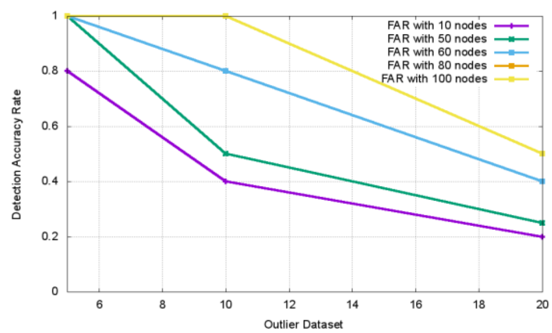


Fig. 7. False Alarm Rate 300 s

In Figure 7, we represent the false alarm rate of our model after a period of 300s. Graphs in Figure 7 show that the model becomes good from 50 sensor nodes.

We have conducted a last simulation that lasted 600s. The Figure 8 represents the results of that simulation without clustering sensor nodes. The DA reaches value 1 with 60, 80, 100 sensor nodes for 5 outliers dataset injected in WSN. For 10 dataset of outliers injected, the value 1 is reached with 100 sensor nodes.

In Figure 9 we represent the results obtained after 600s of simulation while clustering sensor nodes. The DA reaches value 1 with 50, 60, 80, 100 sensor



nodes for 5 outliers dataset injected in WSN. For 10 outliers dataset injected value 1 is reached with 100 sensor nodes.

Fig. 8. Detection accuracy rate after 600 s with clustering

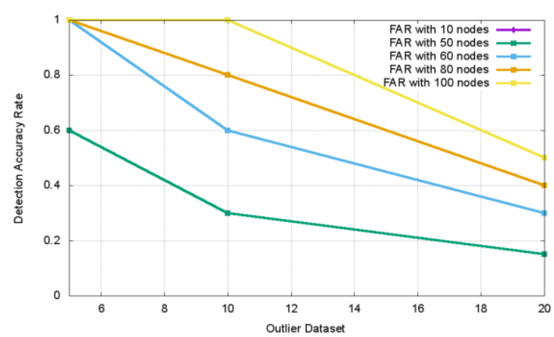


Fig. 9. Detection accuracy rate after 600 s without clustering

Figure 10 shows the false alarm rate of our distributed adaptive model. It shows that our approach becomes good with 50 sensor nodes.

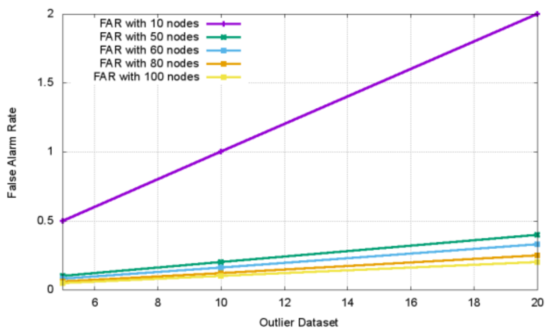


Fig. 10. False Alarm Rate 600 s

Finally, we observed that our distributed adaptive model for outlier detection in WSN achieve good detection accuracy with low false alarm rate. The performance of the model is better when we have a clustered sensor nodes.

6 CONCLUSION

In this paper, we proposed a distributed adaptive scheme for outlier detection in WSNs. In such networks, having significant data among the gathered data by sensor nodes remains a challenge. The proposed approach is based on the Nearest Neighbor Search. We use Mahalanobis distance and Euclidean distance in order to detect outlier while taking into account the size of data gathered by sensor. The proposed approach is suitable for online detection. The performance of the proposed model have been tested by injecting various number of outlier dataset at various times. Experimental results show the effectiveness of our model in term of detection accuracy. The low rate of false alarm indicates the effectiveness of our approach. Moreover, the proposed scheme demonstrates its superiority when sensor nodes are organized in a clustered network. Future work will introduce another objective in the detection, which will consist in the identification of the source of the outlier data.

7 REFERENCES

[1] Aggarwal, S., Singh, J.: Outlier detection using k-mean and hybrid distance technique on multi-dimensional data set. International Journal of Advanced Research in Computer Engineering and Technology 2(9), 2013, 2626–31

[2] Agrawal, S., Agrawal, J.: Survey on anomaly detection using data mining techniques. Procedia Computer Science 60, 2015, 708–713

[3] Ari, A.A.A., Gueroui, A., Labraoui, N., Yenké, B.O.: Concepts and evolution of research in the field of wireless sensor networks. International Journal of Computer Networks & Communications 7(1), 2015, 81–98

[4] Ari, A.A.A., Yenké, B.O., Labraoui, N., Damakoa, I., Gueroui, A.: A power efficient cluster-based routing algorithm for wireless sensor networks: Honeybees swarm intelligence based approach. Journal of Network and Computer Applications 69, 2016, 77–97

[5] Bettencourt, L.M.,Hagberg, A.A., Larkey, L.B.:Separating the wheat from the chaff: practical anomaly detection schemes in ecological applications of distributed sensor networks. In: International Conference on Distributed Computing in Sensor Systems, 2007, pp. 223–239. Springer

[6] Biau, G., Devroye, L.: Lectures on the nearest neighbor method, 2016. Springer

[7] Branch, J.W., Giannella, C., Szymanski, B., Wolff, R., Kargupta, H.: In network outlier detection in wireless sensor networks. Knowledge and information systems 34(1), 2013, 23–54

[8] Chatzigiannakis, V., Papavassiliou, S., Grammatikou, M., Maglaris, B.: Hierarchical anomaly detection in distributed large-scale sensor networks. In: 11th IEEE Symposium on Computers and Communications (ISCC'06), 2006, pp. 761–767. IEEE

[9] Hawkins, D.M.: Identification of outliers, vol. 11, 1980. Springer

[10] Energy Research, Vol. 54, No. 2, 2000, pp. 101-111.

[11] Kamal, S., Ramadan, R.A., Fawzy, E.R.: Smart outlier detection of wireless sensor network. Facta Universitatis, Series: Electronics and Energetics 29(3), 2015, 383–393

[12] Labraoui, N., Gueroui, M., Aliouat, M., Petit, J.: Reactive and adaptive monitoring to secure aggregation in wireless sensor networks. Telecommunication systems 54(1), 2013, 3–17

[13] Labraoui, N., Gueroui, M., Sekhri, L.: On-off attacks mitigation against trust systems in wireless sensor networks. In: Computer Science and Its Applications, 2015, pp. 406–415. Springer

[14] Mahmoud, S.M., Lotfi, A., Langensiepen, C.: User activities outlier detection system using principal component analysis and fuzzy rule-based system. In: Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments, 2012, p. 26. ACM.

[15] Niu, Z., Shi, S., Sun, J., He, X.: A survey of outlier detection methodologies and their applications. In: Inter- national Conference on Artificial Intelligence and Computational Intelligence, 2011, pp. 380–387. Springer
[16] Nourizadeh, S., Deroussent, C., Song, Y.Q., Thomesse, J.P.: Medical and home automation sensor networks for senior citizens telehomecare. In: 2009 IEEE International Conference on Communications Workshops, 2009, pp. 1–5. IEEE.
[17] Rajasegarar, S., Leckie, C., Palaniswami, M., Bezdek, J.C.: Distributed anomaly detection in wireless sensor networks. In: 2006 10th IEEE Singapore International Conference on Communication Systems, 2006, pp. 1–5. IEEE
[18] Salem, O., Guerassimov, A., Mehaoua, A., Marcus, A., Furht, B.: Anomaly detection in medical wireless sensor networks using SVM and linear regression models. In: E-Health and Telemedicine: Concepts, Methodologies, Tools, and Applications, 2016, pp. 466–486. IGI Global
[19] Salem, O., Liu, Y., Mehaoua, A.: Anomaly detection in medical wireless sensor networks. Journal of Computing Science and Engineering 7(4), 2013, 272–284
[20] Shahid, N., Naqvi, I.H., Qaisar, S.B.: Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey. Artificial Intelligence Review 43(2), 2015, 193–228
[21] Sheng, B., Li, Q., Mao, W., Jin, W.: Outlier detection in sensor networks. In: Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing, 2007, pp. 219–228. ACM.
[22] Titouna, C., Aliouat, M., Gueroui, M.: Outlier detection approach using bayes classifiers in wireless sensor networks. Wireless Personal Communications 85(3), 2015, 1009– 1023.
[23] Titouna, C., Aliouat, M., Gueroui, M.: FDS: fault detec- tion scheme for wireless sensor networks. Wireless Personal Communications 86(2), 2016, 549–562
[24] Titouna, C., Gueroui, M., Aliouat, M., Ari, A.A.A., Adouane, A.: Distributed Fault-Tolerant Algorithm for Wireless Sensor Networks. International Journal of Com-munication Networks and Information Security, 9(2), 2017.
[25] Tolle, G., Polastre, J., Szewczyk, R., Culler, D., Turner, N., Tu, K., Burgess, S., Dawson, T., Buonadonna, P., Gay, D., et al.: A macroscope in the redwoods. In: Proceedings of the 3rd international conference on Embedded networked sensor systems, 2005, pp. 51–63. ACM.
[26] Warren, R., Smith, R.F., Cybenko, A.K.: Use of maha- lanobis distance for detecting outliers and outlier clusters in markedly non-normal data: a vehicular traffic example. Tech. rep., DTIC Document, 2011.
[27] Xie, M., Hu, J., Han, S., Chen, H.H.: Scalable hypergrid k-NN-based online anomaly detection in wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems 24(8), 2013, 1661–1670.
[28] Yenké, B.O., Sambo, D.W., Ari, A.A.A., Gueroui, A.: MMEDD: Multithreading model for an efficient data delivery in wireless sensor networks. International Journal of Communication Networks and Information Security 8(3), 2016, 179
[29] Zhang, Y., Hamm, N.A., Meratnia, N., Stein, A., Van de Voort, M., Havinga, P.J.: Statistics-based outlier detection for wireless sensor networks. International Journal of Geographical Information Science 26(8), 2012, 1373–1392
[30] Zhang, Y., Meratnia, N., Havinga, P.: Outlier detection techniques for wireless sensor networks: A survey. IEEE Communications Surveys & Tutorials 12(2), 2010, 159– 170
[31] Ari, A. A. A., Damakoa, I., Gueroui, A., Titouna, C., Labraoui, N., Kaladzavi, G., and Yenke, B. O. 'Bacterial Foraging Optimization Scheme for Mobile Sensing in Wireless Sensor Networks', International Journal of Wireless Information Networks, DOI:10.1007/s10776-017-0359-y, 2017. Springer.

AUTHOR PROFILES:

Dr. Blaise Omer YENKÉ is a Senior Lecturer and researcher in computer engineering. He is actually the head of department of computer engineering at the University Institute of Technology, University of Ngaoundéré, Cameroon. He received his Ph.D. degree in computer science in 2010 from the University of Yaoundé 1 in Cameroon and the University of Grenoble in France, in an international joint supervision. His current research interests include HPC, distributed systems, fault tolerance, sensor networks design and sensor's architecture.

M.Moussa ABOUBAKAR received the B.Sc. degree in mathematics and computer science in 2014 and the M.Sc. degree in computer engineering in 2016 from the Faculty of Science, University of Ngaoundéré, Cameroon. Since September 2016, he is enrolled in a M.Sc. degree in network engineering at Université Paris-Saclay, France. His current research include the improvement of the QoS in wireless networks, Sensors and VANET.

Dr. Chafiq TITOUNA received his Ph.D. degree in computer science in April 2017 from University of Bejaia in Algeria, the Master degree in computer science in 2012 from the University of M'Sila, Algeria. He is currently an Associate Professor at the University of Batna 2. His current research with the Algorithm and Network Team of the LI-PaRAD Lab at the University of Versailles Saint-Quentin-en-Yvelines is focused on Wireless Sensor Networks, 5G and the Cloud Radio Access Network.

Dr. Ado Adamou ABBA ARI received his Ph.D. degree in computer science in 2016 from Université Paris-Saclay in France. He also received the master degree of business administration (MBA) in 2013, the M.Sc. degree in computer engineering in 2012 and the B.Sc. degree in mathematics and computer science in 2010 from the University of Ngaoundéré, Cameroon. His current research with the Algorithm and Network Team of the LI-PaRAD Lab at the University of Versailles Saint-Quentin-en-Yvelines is focused on bio-inspired computing, Wireless Sensor Networks, 5G and the Cloud Radio Access Network.

Dr. Adbelhak GUEROUI is an associate professor in computer engineering at the University of Versailles Saint-Quentin-en-Yvelines (LI-PARAD Lab). He received his HDR (Habilitation de Diriger les Recherches) in 2007 and his Ph.D. degree in computer science both from the University of Versailles Saint-Quentin-en-Yvelines in 2000. His research interests are in the area of wireless networks, especially in QoS routing in broadband networks, Wireless architectures IoT and Cloud Computing. He was involved in several National R&D project in the field of QoS routing and multicasting in constellation of Satellites, ATM broadband core networks, and Wireless IP LAN. Among these projects CORESMA 1996 between the University Paris 6 and THOMSON in which routing algorithms based on adaptive learning has been developed.