Performance of outlier detection techniques based classification in Wireless Sensor Networks

Aya Ayadi, Oussama Ghorbel
CES Research Lab
National Engineers School of Gabes, Gabes University,Tunisia
National Engineers School of Sfax, Sfax University, Tunisia
Email: ing.aya.ayadi@gmail.com ; oussama.ghorbel@gmail.com

M. S. Bensaleh, Abdelfateh Obeid and Mohamed Abid
King Abdulaziz City for Science and Technology, KSA
Digital Research Center (CRNS), Technopark Sfax, Tunisia
Email: mbensaleh@kacst.edu.sa;obeid@kacst.edu.sa
mohamed.abid@enis.rnu.tn

***Abstract*—Nowadays, many wireless sensor networks have been distributed in the real world to collect valuable raw sensed data. The challenge is to extract high-level knowledge from this huge amount of data. However, the identification of outliers can lead to the discovery of useful and meaningful knowledge. In the field of wireless sensor networks, an outlier is defined as a measurement that deviates from the normal behavior of sensed data. Many detection techniques of outliers in WSNs have been extensively studied in the past decade and have focused on classic based algorithms. These techniques identify outlier in the real transaction dataset. This paper aims at providing a structured and comprehensive overview of the existing researches on classification based outlier detection techniques as applicable to WSNs. Thus, we have identified key hypotheses, which are used by these approaches to differentiate between normal and outlier behavior. In addition, this paper tries to provide an easier and a succinct understanding of the classification based techniques. Furthermore, we identified the advantages and disadvantages of different classification based techniques and we presented a comparative guide with useful paradigms for promoting outliers detection research in various WSN applications and suggested further opportunities for future research.**

***Index Terms*—Bayesian networks, Classification based approaches, KPCA, Neural networks, One-class SVM, Outlier detection, Wireless sensor networks.**

I. Introduction

Advances in data processing, electronics and wireless communications have made the vision of wireless sensor nodes an important reality. Wireless sensor networks (WSNs) consist of a large number of these sensor nodes that are networked together. These networks are plenty used and have gained attention in various domains. Most of these applications require precise and accurate data in order to provide reliable information to the end users. However WSNs are vulnerable to intruders and faults due to the low cost nature and harsh deployments of WSNs [1] . Therefore, it is essential to identify potential events as well as erroneous and malicious attacks occurred on the network . Outlier detection methods aim to [illegible] and to refine the collected data and let providing the most useful information to end users. The main challenge is to make a design of an appropriate outlier detection technique designed for WSNs while maintaining the resource consumption of WSNs to a minimum and achieving a high detection rate and low false alarm rate.

Our work discusses detection outlier algorithms based on classification approaches of WSN challenges. The aim of this paper is to help readers better understand the classification based approaches requirements and determine the potential improvements that can be implemented on the existing outliers detection models based on them. Besides, this template introduces a guideline to choose an outlier detection model where effectiveness and efficiency are guaranteed. A comparison of existing detection models is presented and the limitations of each model are mentioned. To the best of our knowledge, there is no recent works that addresses the problem of anomaly detection in WSNs based on the classification approaches. The rest of the paper is organized as follows:

- In Section II, a related works are discussed and the significance of our paper compared to existing works is further highlighted,
- Section III investigates classical outliers detection solutions in WSNs, i.e. those which are based on statistical, clustering, fuzzy logic and Nearest neighbor approaches. We also highlight the drawbacks of these techniques,
- Section IV outlines classification based approaches of outlier detection in WSN and provides a detailed background and preliminaries about outliers detection and related issues about this context in WSNs,
- Section V, we present a comparative guide with useful paradigms for fostering outliers detection research in various WSN applications and suggest further opportunities for future research,
- Finally, Section VI concludes the present work.

II. Related work

Fetching out important information from raw data and identified outliers in it, is a very crucial job. Outlier detection is a broad field, which has been studied in the context of a large number of application domains. Researchers discussed many outlier detection approaches by taking into account various criteria including the detection method such as distribution, density and clustering-based methods. Other Researchers studied some appropriate outliers detections approaches like the nature of data including high-dimensional data, uncertain data and time series data. A methods-based classification of outliers detection models in WSN was described in [2]. In this survey,

978-1-5090-4372-9/17/$31.00 ©2017 IEEE

anomaly detection models were classified into statistical models and non-parametric models based on the techniques used to develop the detection model. Non-parametric techniques were categorized into rule-based, CUSUM-based, data-clustering, density-based, and support vector machines based models. In addition, [3] introduced another comprehensive technique-based classification of outlier detection. In this paper, the non-parametric methods are considered as a part of statistical models. Two additional classes, which are the nearest neighbor-based and spectral-decomposition-based techniques, are added. Bayesian network-based models were classified among classification-based models in addition to the SVM-models. However, Bayesian network models are also classified into three sub-classes, which are the Belief Bayesian, Dynamic Bayesian and Naive Bayesian models. Moreover , a technical taxonomy of detection techniques was presented in [4] and focused on three criteria which are the generality measured by detection accuracy, the speed of detection measured by the computational complexity and the balance between both of them. In 2013, Zhang et al. [5] proposed an ellipsoidal one-class SVM to model the normal behavior of sensor data attributes in WSNs. In addition, authors propose two ellipsoidal one-class SVM-based outlier detection approaches to identify outliers. Thus, they use Spatio-temporal correlation to update the normal behavior of sensor data in a distributed and online manner. This adaptive outlier detection technique achieves the lowest false alarm and the best detection accuracy as compared to existing SVM-based outlier detection techniques designed for WSNs. In the last year, Paola et al. [6] proposes an adaptive distributed Bayesian approach for identifying outliers in collected data by a wireless sensor network. This algorithm improved the accuracy of classification, time complexity and communication complexity, and optimized the metrics for latency and energy consumption if compared to non-adaptive approaches.In addition, Shahid et al. [7] present a taxonomy which is based on the state of- the-art classification scheme about the essential characteristics of the feasibility analysis of outlier detection techniques in harsh environments. Moreover , [8] have proposed a kernel principal component analysis based on Mahalanobis distance for detecting outliers in wireless sensor networks. Mahalanobis distance was used to calculate the mapping of data objects to the feature space so as to separate outliers from normal data. This distance is calculated for the data point to decide whether it is an outlier or normal data. The MD-based KPCA gives better and faster results in finding outliers in wireless sensor networks with low power consumption. In[16], the data is collected from the streets, send it to the datacenter servers and aggregate it to ameliorate certain facilities. Authors compare four anomaly detection techniques to analyze these attacks including Mahalanobis distance, LOF ((local outlier factor), hierarchical clustering and OC-SVM (one-class Support Vector Machines). As a result, they conclude that OC-SVM is the most appropriate technique by achieving the highest true positive rate and the lower false positive rate. This work need to develop the detection performance against anomalies with low consummation. In addition, it is required to develop novel approach to involuntarily arrange the time window size to combine the data collected before fitting the detection model. It is also essential to divide huge networks in areas with a condensed and controllable number of sensors to apply outliers detection technique in a easy way.

III. Limits of classical approaches of outliers detection in WSNs

A. Statistical-Based Approaches

- Parametric techniques are not useful because in most WSN real life applications, there is no prior data distribution knowledge,
- Non-parametric statistical models are not that suitable for real time applications,
- Histograms do not rely on underlying data distribution but they are only efficient for univariate data and cannot find the interactions between attributes in multivariate data,
- Computational cost of handling multivariate data is high.

B. Fuzzy Logic Based Approaches

- It is hard to develop a model from a fuzzy system that requires more fine tuning and simulation before operation,
- Storing the rule-base might require a significant amount of memory. The number of rules grows exponentially to the number of variables,
- Adding spatial and temporal semantics to the decision making process further increases the number of rules,
- Storing a full rule-base on every node might not be reasonable since sensor nodes have limited memory. In addition, constantly traversing a large rule-base might considerably slow down the detection process.
- Reduced, relevant rule subset identification and dynamic rule updating at runtime is a difficult task.

C. Nearest Neighbor-based Approaches

- The computation of the distance between data patterns in multivariate datasets is very expensive,
- The scalability of these models is a major concern,
- Threshold value is used to differentiate outliers from normal object and lower outlierness threshold value will result in high false negative rate for outlier detection,
- Problems arise when a data instance is located between two clusters, the inter-distance between the object of k nearest neighborhood increases when the denominator value increases which leads to high false positive rate,
- Needs to improve the efficiency of density based outlier detection.

D. Clustering Based Approaches

- Dependency on the choice of cluster width in some clustering techniques makes them not suitable for WSN applications,
- Updating the reference model involves a lot of communication overhead and is also computationally expensive,

- Clustering is very computationally expensive with multivariate data because the calculation of the distance measures among all data patterns has high computational cost that make them unsuitable for limited resource devices such as sensors,
- Clustering techniques cannot cope with continuous changes of data streams over time so the normal reference model will be out of date by the time they are used. Although some recent clustering-based models have tackled this issue via incremental learning methods, the computational cost for such methods is too high to be affordable by constrained resource devices.

IV. Performance evaluation of various techniques based classification for outlier detection in WSNs

To suggest additional opportunities for future research, we present a comparative guide with useful paradigms for furthering outliers detection research in various WSN applications. This section aims to evaluate various approaches for outlier detection based on classification in WSNs in terms of accuracy and complexity analysis.

A. General Evaluation Approaches

1) One Class SVM:

- The hypersphere formulation has a better performance and generalization ability for classification as compared to hyperplane formulation. These formulations have a quadratic optimization problem that makes its infeasible for implementation on WSNs, deployed in remote and harsh environments,
- Hyperellipsoid formulations are defined by two parameters. Also, they require the solution of quadratic optimization problems. Whereas, Quarter-sphere and centred ellipsoid are only defined by their radii and they involve a linear optimization problem. These techniques are computationally inexpensive as compared to previous formulations, and more feasible for implementation on WSNs,
- Hyperplane, hypersphere and quarter-sphere based formulations use a Euclidean distance of data samples from the center of distribution to detect outliers. This distance measure is highly dependent on the difference between the absolute values of attributes. For that, it assumes that data are equally distributed in all directions.
- However, hyperellipsoid and centred hyperellipsoid use Mahalanobis distance to detect outliers. This distance measure denotes the deviation of a data sample from the center of distribution without considering the spherical nature of data distribution. Many researches prove that classifiers based on Mahalanobis distance reach better generalization as compared to those formulations based on Euclidean distance.
- The hypersphere formulation has a better performance in terms of classification as compared to hyperplane. Furthermore, despite the fact that the quarter-sphere formulation has the same performance as that of hypersphere formulation, its computational complexity is less than hypersphere formulation. In addition, the hyperellipsoid formulation classifies all data points correctly by fitting a tight ellipsoid on the data distribution. All the data vectors, which are away from the data distribution, are classified as abnormal. Though the centered ellipsoid formulation has a similar performance as hyperellipsoid, it has a reduced computational complexity.

2) Bayesian Networks:

- Nave Bayes classifiers require more effort in the data recording phase because it requires labeled training data to clearly recognize the defined activities,
- During the training phase, the classifier requires significant computational overheads to learn a probabilistic graphical model. However, this model may have to be learned in time due to the dynamic nature of data distributions in harsh environments,
- Bayes Classifiers are offline and not adaptive. Thus, a big amount of training data is required to learn a model. In addition, the training phase is repeated for each distributions change,
- Naive bayes assumption can be used to reduce the complexity of the learning phase. However, this assumption also leads to a loss of correlations between various attributes ultimately leading to a reduction in the generalization ability.

3) KPCA:

- Some current KPCA reconstruction methods equally weigh all the features; it is impossible to weigh the importance of some features over the others,
- Some other works only consider robustness of the principal subspace; they do not address robust fitting. However, this approach does not handle intra-sample outliers,
- KPCA is computationally intensive and takes a lot more time compared to PCA. The reason behind this is the that the number of training data points in KPCA is much higher than PCA. Therefore, the number of principle components that need to be estimated is also much larger,
- Both the RE and OCSVM produce a decision boundary that is overly broad. Thus, it does not satisfactorily fit the normal data because many potential outliers would not be detected. However, the MD induced boundary seems to capture much better the overall structure of the normal data.
- KPCA using the Mahalanobis distance (MD), as shown in Fig 4, is more sensitive to the detection of FPR and DR than KPCA using reconstruction error (RE) and OCSVM.

4) Neural Networks:

- Multi-class classification based techniques rely on the availability of accurate labels for various normal classes, which is often not possible.
- Assigning label to each test instance becomes a disadvantage when a meaningful outlier score is desired for

the test instances being subject to classification based techniques. However, some classification techniques that obtain a probabilistic score from the output of a classifier can be used to address this issue.

- MLPs do not learn closed class boundaries. Thus, they tend to generalize and assign a novel instance to one of the learnt classes. A modification to the MLP structure is proposed to fix the problem of learning the boundaries well. This is done through constructing hyperplanes between different classes to better separate them from each other.

B. Accuracy Analysis

The data vectors which have ellipsoidal distribution, such as ellipsoidal SVM-based online outlier detection technique (EOOD)[9] and Ellipsoidal SVM-based adaptive outlier detection technique (EAOD)[9], perform better than spherical SVM-based batch outlier detection technique and spherical SVM-based adaptive outlier detection technique. Ellipsoidal distribution formulations take into account the correlation of data attributes and have the better understanding of multivariate nature of data distribution. On the other hand, spherical SVM based outlier detection techniques ignore this criterion of correlation between data attributes and use a spherical boundary to fit the data. Therefore, these kinds of methods have a low detection rate and a high false alarm rate in case of non-spherical data distribution. In addition, ellipsoidal SVM-based techniques perform better than ellipsoidal SVM-based batch outlier detection technique (EBOD)[10][11][12] due to the fact that they exchange essential ellipsoidal information with neighbouring nodes for reliable outlier detection. However, spherical SVM-based techniques assume that data vectors are distributed around the center of the mass in an ideal spherical shape. Concerning these data vectors such as spherical SVM-based adaptive outlier detection technique (SAOD)[9] and spherical SVM-based batch outlier detection technique (SBOD)[10][11][12], they achieve better detection accuracy and lower false alarm compared with ellipsoidal SVM-based such as EOOD as well as EAOD and EBOD. While ellipsoidal SVM-based approaches take into account the correlation of data attribute, they as well as spherical SVM-based techniques do not perform spherical data distribution. Moreover, among spherical SVM-based techniques, SAOD performs better than SBOD. SAOD alleviates the influence of outliers by using median and median absolute deviation (MAD). Besides, this technique exchanges this spherical information with neighbouring nodes for reliable outlier detection. Regarding the real data, EAOD achieves the lowest false alarm and the best accuracy. This technique considers the correlation of data attributes as well as ellipsoidal information (median, covariance matrix) from neighbouring nodes. As a result, EADO has the highest detection accuracy. On the contrary, EOOD does not update the normal profile, while the data distribution has changed. As a result, this technique archives the highest false alarm. Therefore, EAOD generates the best performance compared to other SVM-based techniques. Thence, using correlation among data attributes and understanding data distribution are essential to model the normal behavior of data vectors. These assumptions design a suitable outlier detection technique.

For methods based KPCA, Mahalanobis kernel has been used recently in the field of WSN. Specially based outlier detection was introduced in several works. Kernel PCA performance was showcased in comparison to other established kernel-based methods [13]. To compute the Kernel PCA transformation of a set of test patterns, this approach chooses a training set and a suitable projection dimensionality p, and finally, computes the Mahalanobis distance (MD) for each of these test patterns. Given the projection dimensionality p, outliers are identified as data points, whose MD exceeds an appropriately established threshold value MaxDist. Same researches like [8] discussed that kind of outlier detection techniques and results have shown that using Mahalanobis distance is more beneficial to detect outliers. Then, this comparison reveals that both the RE and OCSVM may not be an effective measure of deviation from normalcy when compared to using the MD. In addition, RE produces a decision boundary that is overly broad. Thus, it does not satisfactorily fit the normal data because many potential outliers would not be detected. Consequently, the MD based method has an important advantage compared to the RE-based method and OCSVM that perfectly detects the outliers. This technique presented in [13] is more sensitive to the detection of FPR and DR than KPCA using reconstruction error (RE) and OCSVM. However, the MD induced boundary seems to capture much better the overall structure of the normal data.

C. Complexity analysis

A summary of the various complexities involved in different schemes is provided in table I [9] [12].

TABLE I
COMPLEXITY ANALYSIS OF VARIOUS DETECTION SCHEMES

Techniques	Computational complexity (per node)	Memory Complexity (per node)	Communication complexity (per link)
Distributed QSSVM	O(n^2)	O(np+n)	O(l)
Centralised QSSVM	O(s^2n^2)	O(snp)	O(np)
Centralised CESVM	O(n^2 +m^2 n)	O(np+mn)	O(np)
EBOD	-	O(kdn^2))	O(n)
SBOD	O(n)	O(kdp)	O(dn+d^2)
EOOD	O(n^2)	O(dn^2)	O(dn)
EAOD	O(n^2)	O(dn^2)	O(dn)

The CESVM scheme involves the computation of a kernel matrix k (also called a Gram matrix) with a computational complexity of O (n^2), an eigen-decomposition of k, and solving a linear optimization problem. Eigen-decomposition requires a complexity of O (n^3). However, Williams et al. [14] have observed that the eigen values decay exponentially for a wide variety of kernels such as the RBF. Hence, a reduced rank representation of the Gram matrix can be obtained by exploiting this observation. The incomplete Cholesky decomposition method [15] achieves this with complexity O(m^2n),

where m¡¡n is the low rank approximation of the Gram matrix k. Therefore, the total computational complexity involved excluding the complexity in linear optimization is O(n^2+m^2n). CESVM uses linear programming optimization. There are many algorithms available for the linear programming in the literature. The simplex algorithm is efficient, but it has been shown to have worst-case exponential complexity in the number of variables [10]. Polynomial time algorithms such as the interior point methods incur O(n^3) arithmetic operations and have a complexity of O(L) iterations [14], which is the number of variables and the size of the optimization problem, i.e., roughly the number of bits required to represent the problem. Once the linear optimization is performed, each node has to keep the eigenvalues and the eigenvectors P in memory. Thus, the memory complexity is O(mn++np), which includes the memory complexity required to keep data vectors in memory.

On the other hand, The QSSVM involves computation of a kernel matrix K and solving a linear optimization problem. The complexity of computing the matrix k of size nn is O(n^2) [14]. QSSVM incurs a computational complexity of linear programming as explained in CESVM. In the distributed scheme, once optimization is performed, each node keeps only the radius value and the norms of the data vectors in memory in addition to the data vectors. Hence, the memory complexity for each node is O (np+n). On receipt of the radius information from the children, the parent node computes the global radius using one of the techniques such as median as well as maximum and minimum. As a result, this strategy involves a computational complexity of O (l), where l is the number of children of the parent node Sp. When the sensor nodes receive the global radius from their parent node, they compare the norms of their local data vectors to the global radius. Then, this involves a single pass over the data vectors with computational complexity O (n). In contrast, centralized detection involves the communication of all data measurements to a central node with a communication complexity of O (np) per link giving a total of O (snp), where p is the dimension of data vectors and s is the number of sensor nodes in the network. QSSVM is run at the central node on the collected data with a maximum computational complexity of O (s^2n^2). The memory complexity at the central node is O (snp).

However, the communication complexity of the ellipsoidal SVM-based distributed techniques depends on the transmission parameters such as local hyper-ellipsoid radius information as well as the median and covariance matrix. The communication overhead in EOOD for each node is O(n^2), where n is the dimension of observations. Each node transmits only information about its local hyper-ellipsoid radius information, the median, and covariance matrix once at the initial training phase. EAOD dont update the radius information during online outlier detection that only possibly communicates the covariance matrix and the updated median as well as the radius information with nodes at the end of a sliding time window.

The computational complexity in EOOD and EAOD is related to the computation of some parameters such as the median and the covariance matrix as well as the distance between every new observation and the origin, thus, the linear optimization function. The computational complexity of these techniques mainly depends on solving a linear optimization problem, which is represented as O (h). In addition, the complexity of computing covariance matrix, is represented as O (dn^2) where d is the number of new observations to be classified. However, EBOD also computes kernel matrix and the transformation of a centred kernel function. This complexity is represented by O (k). Accordingly, the maximum computational complexity of EBOD for each node is O (kmd^2).

The memory complexity of EOOD and EAOD is generally related to keeping the observations of the size of sliding window in memory. This complexity is represented as O (dn), where d is the dimension of observations and n is the number of new observations. Storing other parameters such as covariance matrix with a complexity of O(n^2) is negligible since d ¿ n. Due to the fact that EBOD needs to keep m*m kernel function, its memory complexity of each node is O(dn+n^2).

Moreover, similar researches like [8] have shown that for the computational complexity using KPCA based Mahalanobis kernel, only the testing phase should be considered which is conducted online. In this phase, the calculation of the reconstruction error measure based on Mahalanobis distance for each new observation is done by dividing its projection on the PC space. This is carried out through corresponding eigenvectors which are already calculated in the training phase. The upper bound computational complexity involved in this process is O (m), where M is the number of observed variables. The training phase which involves the calculation of the PCs has a time complexity of O (nm^2) where N and M are the sizes and the dimensions of the training set respectively. For the training phase of the OCSVM, the relative complexity is O (n^3) as it needs to solve an optimization problem to compute the hyper plan that separates normal and outlier data. The complexity of online testing phase in the theme of our model structure is O(m) where m is the number of observed variables which is equivalent to the complexity incurred by our proposed model. The retraining of the OCSVM will cause a high power consumption which makes it unsuitable for anomaly detection compared to a model using KPCA based Mahalanobis kernel.

V. Experiment and Discussion

The experiment was performed on Digital Research Center Technopark Sfax (CRNS). The data samples were gathered from a deployment of WSN in static and dynamic environments to detect leak in water pipeline. The database contains 694 samples of 2 numeral classes (network with leak and network without leak). Here, our experiment was performed based on set of pressure. The pressure variation curve is represented by the Figure 1.

In the experiment, we chose the last 147 samples of each class for testing, the remaining 200 samples for training. Consequently, the total number of testing samples is 494 while

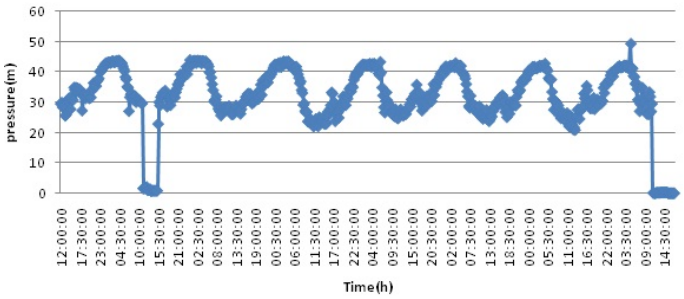

Fig. 1. Pressure variation curve

the total number of training samples is 400. In addition, the samples of each group are labeled accordingly. In the present study, the effective data were trained and tested using five different classifiers: Nave Bayesian network, Neural Network, K^{th} Nearest Neighbors, Support Vector Machine and Decision Tree. To calculate the precision of different techniques, we calculate the detection rate that means how many outliers are correctly detected, and false positive rate (FPR) that indicates how many normal data are incorrectly detected as abnormal data.The effectiveness outlier detection technique should reach a low FPR and high DR. The performance comparison of the classifiers has been presented in table II.

TABLE II
EXPERIMENTAL RESULTS FOR VARIOUS CLASSIFIERS

	DR	FPR
Naive Bayesian Network	88%	12%
Neural network	92%	8%
SVM	93%	7%
Decision tree	87%	13%
K^{th}Nearest Neighbors	81%	19%

In our binary classification problem (network with leak or without leak), Support Vector machine outperforms Naive Bayesian Network, Decision Tree, K^{th} Nearest Neighbors and a very small difference with Neural Network. . If we look into the reason for this performance, we find that SVM is based on the hyper-plane boundary and can find the optimal compromise between the complexity of the model and the learning ability to obtain the best generalization. The average detection accuracy in detection leakages is 93 % for SVM, 92 % for neural network, 88 % for Naive Bayes , 87% for Decision tree and 81% for K^{th} Nearest Neighbors. We can't make a conclusion that a classification technique is better than another because it may work well in a certain data constraints, but worse in others. By means of large data sets or a combination of these classifiers, the performance can be improved supplementary.

VI. CONCLUSION

This paper gives a brief overview about the classification strategies for outlier detection techniques in WSNs and discusses the feasibility of various types of classic techniques for WSNs. The aim of this template is to present a detailed analysis of various technique based classification formulations for outlier detection in Wireless sensor networks deployed. These formulations include one-class SVM, Bayesian networks, neural networks and KPCA. We have provided an easier and more succinct understanding of this techniques and we presented a comparative analysis that has been carried out in terms of characteristics like accuracy, communication, computational and memory complexity to identify the improvement and feasibility of various techniques presented.

ACKNOWLEDGMENT

This work was supported by King Abdulaziz City for Science and Technology (KACST) and the digital Research Center of Sfax (CRNS) under a research grant (project no. 35/1012).

REFERENCES

[1] I.F. Akyildiz, T. Melodia, Kaushik R. Chowdhury. *A survey on wireless multimedia sensor networks*. Journal Computer Networks: The International Journal of Computer and Telecommunications Networking, Volume 51, Issue 4, Inc . New York, NY, USA, United State, 2007.

[2] Rajasegarar, S.; Leckie, C.; Palaniswami, M. *Detecting data anomalies in wireless sensor networks*. Secur. AdHoc. Sens. Netw. 2009, 3, 231259.

[3] Y. Zhang, M. Nirvana, H. Paul *Outlier Detection Techniques For Wireless Sensor Networks: A Survey*, University of Twente, P.O.Box 217 7500AE, Enschede, The Netherlands, 2010.

[4] Xie, M.; Han, S.; Tian, B.; Parvin, S. *Anomaly detection in wireless sensor networks: A survey* . J. Netw. Comput. Applic. 2011, 34, 13021325.

[5] Yang Zhang, Nirvana Meratnia, Paul J.M. Havinga. *Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine*.Ad Hoc Networks, Volume 11, Issue 3, May 2013, Pages 1062-1074

[6] A. De Paola, S. Gaglio, G. Re, F. Milazzo and M. Ortolani, *Adaptive distributed outlier detection for WSNs*,IEEE Trans. Cybern., vol. 45, no. 5, pp. 888-899, 2015

[7] Shahid, N., Naqvi, I. H., and Qaisar, S. B , *Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey*,Artificial Intelligence Review, 43(2), 193-228.(2015)

[8] O. Ghorbel, Walid Ayedi, Hichem Snoussi, and Mohamed Abid, *Fast and Efficient Outlier Detection Method in Wireless Sensor Networks*, IEEE journal, June 2015

[9] Y. Zhang, N. Meratnia, Paul J.M. Havinga, *Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine*, Ad Hoc Networks, Volume 11, Issue 3, May 2013, Pages 1062-1074

[10] S. Rajasegarar, C. Leckie, M. Palaniswami,, J. C. Bezdek *Quarter sphere based distributed anomaly detection in wireless sensor networks* , Proceedings of IEEE International Conference on Communications, pp. 3864-3869, 2007.

[11] S. Rajasegarar, C. Leckie, M. Palaniswami, *CESVM: Centered hyperellipsoidal support vector machine based anomaly detection*, in: Proceedings of IEEE International Conference on Communications, Beijing, China, 2008, pp. 16101614.

[12] S. Rajasegarar , C. Leckie , J. C. Bezdek and M. Palaniswami, *Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks*, IEEE Trans. Inf. Forensics Security, vol. 5, no. 3, pp. 518-533, 2010

[13] D, Mingtao, T, Zheng and X, Haixia *Adaptive kernel principal component analysis*, Signal Process, pp 1542-1553, 2010.

[14] C. K. I. Williams and M. Seeger *The effect of the input density distribution on kernel-based classifiers*, in Proc. 17th Int. Conf. Machine Learning, San Francisco, CA, 2000, pp. 11591166.

[15] F. R. Bach and M. I. Jordan *Kernel independent component analysis*, J. Mach. Learning Res. (JMLR), vol. 3, pp. 148, 2003.

[16] Garcia-Font, V., Garrigues, C., and Rif-Pous, H. *A Comparative study of anomaly detection techniques for smart city wireless sensor networks*,Sensors, 16(6), 868 (2016).