Journal of Computer Science

Original Research Paper

A Clustering based Approach for Contextual Anomaly Detection in Internet of Things

Dina ElMenshawy, Waleed Helmy and Neamat El-Tazi

Department of Information Systems, Faculty of Computers and Information, Cairo University, Cairo, Egypt

Article history
Received: 23-05-2019
Revised: 11-07-2019
Accepted: 23-08-2019

Corresponding author:
Dina ElMenshawy
Department of Information Systems, Faculty of Computers and Information, Cairo University, Cairo, Egypt
Email: d.ezzat@fci-cu.edu.eg

Abstract: Internet of Things (IoT) is a network which connects different communication devices with the internet to attain quick, robust and real-time information transfer and communication, achieving intelligent management. IoT is still in its infancy so it faces numerous challenges varying from data management to security concerns. Sensors generate enormous quantities of data that need to be handled efficiently to have successful deployment of IoT applications. Concerning data management, a great challenge that faces the IoT environment is the detection of contextual anomalies. Contextual anomaly detection is a sophisticated task because the context has to be taken into consideration in the anomaly detection process rather than checking only the deviation of the data value as in point anomaly detection. As a result, in this paper, a novel clustering based algorithm is proposed to detect contextual anomalies in Internet of Things. Attributes were separated into two different categories, namely contextual attributes and behavioral attributes. K-Means clustering technique was applied on the contextual and behavioral attributes separately, then the intersection between the contextual and behavioral clusters was used to detect the contextual anomalies. Moreover, the algorithm was applied on a real room occupation dataset of size around 20,000 records and the experiments showed promising results.

Keywords: Internet of Things, Contextual Anomaly, Clustering

I. Introduction

Internet of Things (IoT) is a network which links various communication devices with the internet to achieve fast, reliable and real-time information transfer and communication, achieving intelligent management. IoT gains popularity through the few recent years. The IoT paradigm evolves as a result of the huge availability of the internet access, communication protocols and sensors. It utilizes current technologies, such as: Cloud and mobile technologies (Yue *et al*., 2015).

IoT comprises a group of things which are enclosed with electronics, software, sensors and network connectivity that allow these things to capture and exchange data. A thing denotes an object in the real world which has an identity and can be embedded into a communication network. As a result, objects can be managed remotely, permitting integration between the physical and virtual worlds (Elbouanani *et al*., 2015).

The major characteristics of IoT are:

(1) **Comprehensive perception:** Sensors capture data anytime and anyplace.
(2) **Reliable transmission:** Objects access information networks and achieve robust information interaction
(3) **Intelligent processing:** Analyze the huge quantities of sensors' data by using various technologies and attain intelligent decision making and control (Elbouanani *et al*., 2015; Ray *et al*., 2016)

The IoT paradigm allows for:

(1) Device to device information communication
(2) Device to people information communication
(3) Device to environment information communication by the consolidation of information space and physical space (Ray *et al*., 2016)

IoT consists of three main layers as shown in Fig. 1.

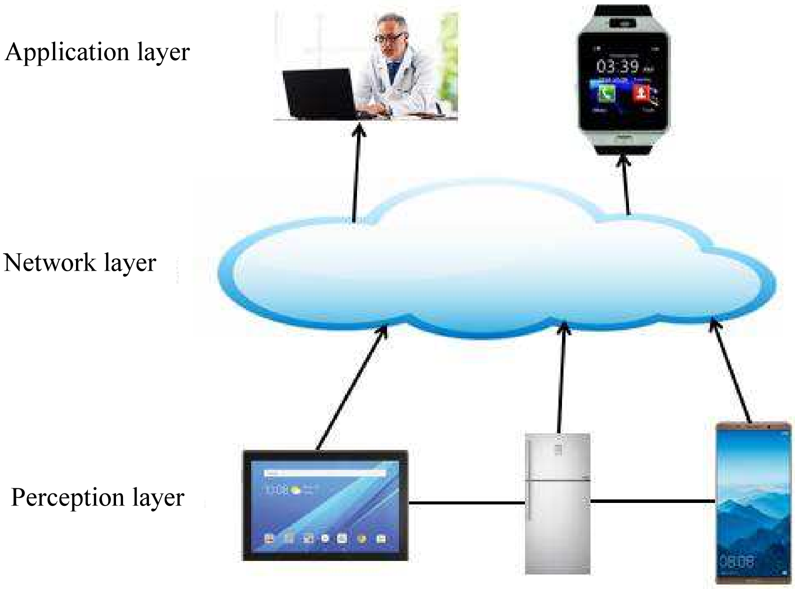

Fig. 1: IoT architecture (adapted from (Rose, 2014))

IoT layers are described as follows:

(1) **Application layer:** Denotes the application service support system
(2) **Network layer:** Includes the communication network infrastructure
(3) **Perception layer:** Comprises the sensor based devices and environmental objects. The acquired data from this layer is transmitted to the network layer for processing and analysis (Rose, 2014; Kraijak and Tuwanut, 2015)

II. Motivation

IoT is still in its beginning so it faces a lot of challenges ranging from data management issues to security concerns. Security threats are huge in the IoT environment because of the physical accessibility to the objects, which are embedded with the sensors. Also, the majority of the objects communicate wirelessly so the environment is highly susceptible to security attacks (Nalbandian, 2015). As a result, a lot of research and investigation are needed to deal with these issues to increase the IoT usage in various applications.

Regarding data management issues, numerous challenges arise since sensors produce enormous amounts of data that need to be managed efficiently. Extracting meaningful information from the massive amounts of data is not an easy task. A huge challenge that faces the IoT paradigm is detecting anomalies of sensors' data. An **anomaly/outlier** refers to a data point which greatly varies from the remaining data points, as though it was induced by another approach (Han *et al*., 2012).

There are three types of anomalies, described as follows:

(1) **Global/point anomaly:** In a specific dataset, a data point is considered a global anomaly if it varies greatly from the other data points (Han *et al*., 2012)
(2) **Contextual anomaly:** In a certain dataset, a data point is considered a contextual anomaly if it notably varies in the defined context. Contextual anomalies are also denoted as conditional anomalies because they depend on a specific context. Consequently, to detect contextual anomalies, the context has to be specified as a major component of the problem definition. In contextual anomaly detection, the attributes under consideration are divided into two types:

- **Contextual attributes:** These features define the object's context. Context can refer to a time interval, location, weather, country, etc.
- **Behavioral attributes:** These features represent the object's characteristics and are used to determine whether a data instance is an anomaly in the context which it belongs (Han *et al*., 2012)

(3) **Collective anomaly:** In a certain data set, a subset of data points forms a collective anomaly if the points entirely vary greatly from the rest of the dataset. On the other hand, the individual data points may not be considered anomalies (Han *et al*., 2012)

Data anomaly detection in IoT is considered a complicated task because:

(1) The notion of anomaly in IoT is domain dependent
(2) There is a high noise rate in the data, since sometimes sensors have low quality and power, so noisy data may be interpreted as anomalies and vice versa (Nalbandian, 2015)

In IoT, anomaly detection has several challenges in general and in contextual anomaly detection in particular, because, the context has to be taken into account in the anomaly detection process.

In this paper, we explore the problem of contextual anomaly detection in IoT. A clustering based approach is proposed to detect contextual anomalies in the IoT environment. K-Means (Han *et al*., 2012) clustering technique was used as the clustering technique.

The rest of this paper is organized as follows: Section 3 presents the literature work. Section 4 presents the proposed algorithm along with the experiments and results. Section 5 concludes the paper.

III. Literature Work

We noted that contextual anomaly detection is not extensively explored in the literature. Few literature work explored the problem of contextual anomaly detection in general and in IoT in particular. In this section, existing research related to using contextual information in the anomaly detection process is presented.

Kosek (2016), the authors presented a contextual anomaly detection technique to detect malicious voltage control actions in the low voltage distribution grid. The anomaly detection technique applied artificial neural networks to define a distributed energy resource's behavior under control. The detection system examined the distributed energy resource's behavior, control actions and power system impact and was tested with an ongoing voltage control attack in a simulation environment. The results of applying the detection system on a real photovoltaic rooftop power plant data showed that the contextual anomaly detection gave better performance than the point anomaly detection.

Liu *et al*. (2017), an unsupervised anomaly detection technique that does not need prior knowledge in discovering anomalous events was proposed. Anomalies were defined as groups of anomalous objects varying contextually from their spatial and temporal neighbors. The proposed algorithm was effective in filtering out noisy pixels, detecting spatial-temporal anomalies and grouping those anomalies into anomalous events.

Hayes and Capretz (2014), a novel approach for contextual anomaly detection in big sensor data was presented. The point anomalies were further processed by a context aware anomaly detection algorithm using a clustering technique to decide whether the anomaly was contextually anomalous. The proposed algorithm provided improvements over point anomaly detection algorithms.

Radon *et al*. (2015), a novel anomalous detection framework that used contextual information to reduce false alarms through contextual verification was proposed. Contextual features were extracted from the domain knowledge. Those features were considered the factors which determine whether the points were anomalous or not.

Thah and Sitanggang (2016), the authors presented an approach to discover contextual anomalies on hotspot data based on climate context, i.e., rainfall. Contextual anomalies were detected using the results of clustering on the daily hotspot frequency attribute and rainfall attribute. K-Means algorithm was applied to detect the contextual anomalies. The contextual anomalies were the hotspots that have high daily frequency with high rainfall.

Leach *et al*. (2014), the paper explored the problem of discovering human behavioral anomalies in crowded surveillance environments. The proposed algorithm focused on detecting subtle anomalies in a behaviorally heterogeneous surveillance scene. A novel unsupervised context-aware technique was implemented. Social context and scene context were used to enhance the behavioral analysis. It was found that in a crowded scene the usage of mutual information based social context allowed the capability to inhibit self-justifying groups and disseminate anomalies in a social network, leading to a better anomaly detection capability.

Berrocal *et al*. (2017), the authors introduced a monitoring system which utilized the contextual information to find cognitively impaired person's routines and deviations. Routines were extracted though monitoring the daily life of the cognitively impaired elderly. The authors presented a system which used the daily routine information to detect variations from the normal behavior.

IV. Anomaly Detection

A. Anomaly Detection Techniques

If class labels of normal and anomalies are available, anomaly detection models can be constructed. The techniques of anomaly detection are categorized into three types:

(1) Supervised methods
(2) Semi-supervised methods
(3) Unsupervised methods

The normality and abnormality of data points can be modeled though the supervised methods. Domain experts analyze the dataset and label the underlying data. Then, the anomaly detection process can be represented as a classification problem. The classifier is learnt to detect anomalies and the sample is used for training and testing.

On the other hand, if labels of data points are not available, an unsupervised learning method has to be used which makes an assumption: The normal data points are somewhat clustered. Normal points do not have to exist in one group sharing high similarity, but they can exist in several groups, where each group has unique characteristics. A point is considered anomaly if it exists far away from any of these categories of normal points.

In semi-supervised methods, some labeled examples are available but the number of these labeled examples is usually small. Only a little set of the normal and/or anomalies have labels, but the majority of the data points are unlabeled (Han *et al*., 2012; Hand *et al*., 2001).

B. Contextual Anomaly Detection

Contextual anomaly detection is considered a problematic issue compared to the point anomaly detection. This is because the context has to be taken into consideration to determine whether the point is anomaly. On the other hand, in point anomaly detection, only the point is considered anomalous if it has a great deviation in comparison to the rest of the dataset. In contextual anomaly detection, the same value of the data instance can be anomalous in a specific context while normal in another. For example, a specific value of the temperature can be normal in summer in a certain county while anomalous in winter (Kosek, 2016). Also, the value of the flow of cars in a certain time can be normal (for example in rush hours) while it will be considered anomalous in another time (in midnight).

As a result, to detect contextual anomalies, the context has to be defined as a part of the research problem. One way to put the context into consideration is to add the contextual attributes to the behavioral attributes in the anomaly detection process. Instead of including the behavioral attributes as a feature vector in the anomaly detection process, the contextual attributes will be added to the feature vector, i.e., the problem will be dealt as a point anomaly detection process (Kosek, 2016).

Another way to consider the domain context, is to divide the dataset into contextual groups and examine the behavioral attributes for every context separately, leading to several point anomaly detections processes (Kosek, 2016).

In this research, we propose a novel way of detecting contextual anomalies based on a clustering technique.

C. Technique Used

The conception of anomalies is highly associated to that of clusters. Clustering based techniques discover anomalies by analyzing the relationship between data points and clusters. Many clustering techniques can be used in unsupervised anomaly detection. The main idea is to discover clusters first, then the data points which do not belong to any clusters or belong to small clusters are considered anomalies (Han *et al*., 2012). Since class labels in most datasets are not available so unsupervised techniques are usually used in various applications.

In our proposed algorithm, K-Means clustering technique was used to cluster the contextual and behavioral attributes. We chose K-Means technique because it has high scalability in processing huge datasets.

K-Means is a centroid based technique which uses the centroid of a cluster to represent that cluster. The cluster's centroid represents its center point. The centroid can be calculated in different ways such as by the mean or medoid of the points allocated to the cluster. The difference between a data point and the centroid of the cluster, is calculated by the Euclidean distance.

An objective function is utilized to evaluate the partitioning quality so that points within a cluster are similar to each other but dissimilar to points in other clusters. The goals of the K-Means clustering technique are achieving high intracluster similarity and low intercluster similarity (Han *et al*., 2012; Hand *et al*., 2001).

D. Proposed Algorithm

The main steps of the proposed algorithm are as follows:

(1) Divide the attributes into two categories: Contextual attributes and behavioral attributes
(2) Apply K-Means clustering technique on the contextual attributes
(3) Apply K-Means clustering technique on the behavioral attributes
(4) Get the data tuples resulting from the intersection between each contextual cluster with all behavioral clusters as depicted in Fig. 2. The intersection criteria is that the values of the contextual and behavioral attributes which exist in the same data tuple

E. Optimal Number of Clusters

In order to determine the optimal number of clusters (k), two techniques were used, namely: The elbow method and silhouette method (Oh and Kim, 2017).

1) Elbow Method

Elbow method can be described as follows:

- Start with k = 2 and continue increasing it in every iteration by 1
- Calculate the cost that results with the clustering
- Monitor the evolution of the within-cluster sum of squares
- Check the elbow point (Oh and Kim, 2017)

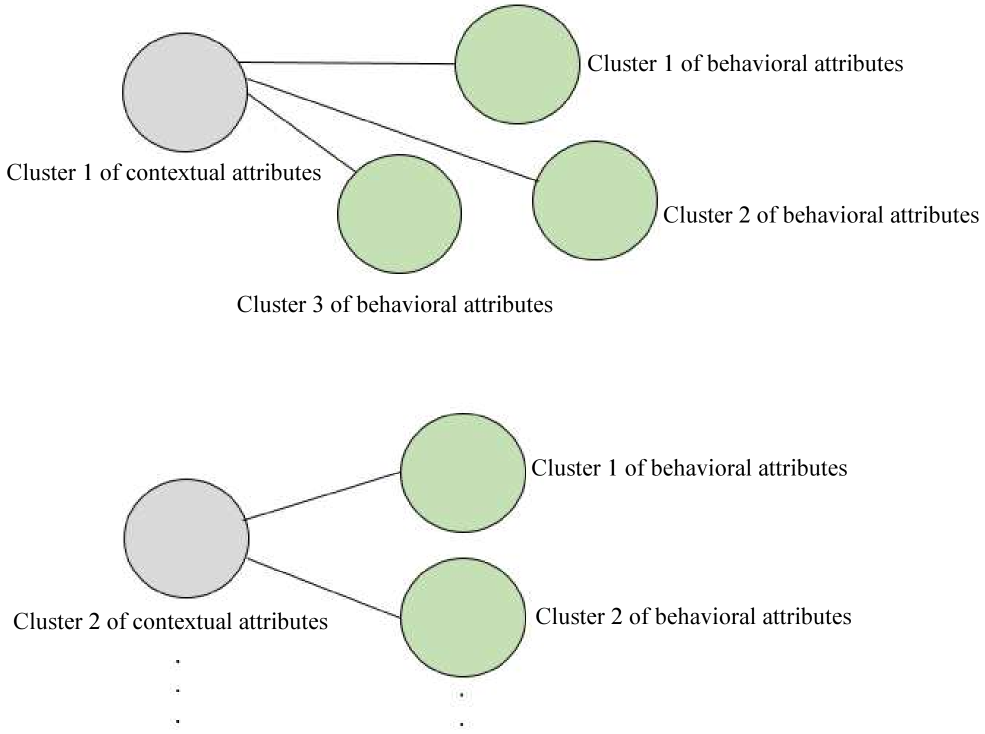
Cluster 1 of behavioral attributes
Cluster 1 of contextual attributes
Cluster 2 of behavioral attributes
Cluster 3 of behavioral attributes
Cluster 1 of behavioral attributes
Cluster 2 of contextual attributes
Cluster 2 of behavioral attributes

Fig. 2: Intersection of contextual and behavioral clusters

The elbow occurs at the point where adding an additional cluster does not decrease remarkably within-cluster sum of squares. If the number of clusters increases, the average distortion will diminish, each cluster will have fewer points and the points will be closer to their corresponding centroids. On the other hand, the enhancement in average distortion will decrease as the value of k increases (Oh and Kim, 2017).

The value of k at which the enhancement in distortion decreases the most is named the elbow, the point at which we should terminate splitting the data into more clusters. The elbow method shows the amount of the cost function produced by various values of k. Usually, the first clusters will enhance the clustering result, but at a certain point the additional gain will diminish greatly and produce an angle in the curve (Arroyo *et al*., 2017).

2) Silhouette Method

The second technique used to determine the most suitable number of clusters is the silhouette method. The silhouette method is a measure of the similarity of an object to its cluster compared to other clusters. It ranges from -1 to 1, where a large value means that the object is highly matched to its own cluster and low matched to the other clusters (Wang *et al*., 2018). If the majority of objects have high values, then the clustering result is considered successful. If a lot points have low or negative values, so the clustering result is not considered appropriate, since it led to either having a lot of clusters or few clusters. The silhouette value can be computed with any distance measure, like the Euclidean distance or the Manhattan distance (Meng *et al*., 2018; Rajeswari *et al*., 2018).

F. Dataset

The dataset used includes data related to detection of room occupancy from temperature, humidity, light and CO_2 (Candanedo and Feldheim, 2016). Ground-truth occupancy was acquired from time stamped pictures that were captured every minute. The dataset has the following attributes:

1. Date and time (day-month-year-hour-minute-second)
2. Temperature, in Celsius
3. Relative Humidity, %
4. Light, in Lux
5. CO_2, in ppm
6. Humidity Ratio, derived measure from temperature and relative humidity, in kg-water-vapor/kg-air (Candanedo and Feldheim, 2016)

A snapshot of the dataset is shown in Table 1.

The proposed algorithm was applied on the dataset consisting of around 20,000 records. Attributes were divided into two categories: Contextual attributes and behavioral attributes.

Table 1: Snapshot of the dataset

Date	Time	Temperature	Relative humidity	Light	CO_2	Humidity ratio
2/2/2015	2:19:01	23.73	26.27	585.2	749.2	0.004764
2/2/2015	2:19:02	23.72	26.29	578.4	760.4	0.004772
2/2/2015	2:19:03	23.7	26.23	572.6	769.6	0.004765

Table 2: Intersection between contextual and behavorial clusters

Contextual cluster number	Behavioral cluster number	Number of intersecting instances
CC1	BC5	692
	BC8	245
	BC10	4
CC2	BC13	3
	BC15	6166
CC3	BC6	606
CC4	BC4	585
	BC11	306
CC5	BC1	1200
CC6	BC3	281
CC7	BC6	210
CC8	BC13	469
CC9	BC1	7
	BC10	1675
	BC14	1
CC10	BC7	1355
	BC11	114
CC11	BC9	416
CC12	BC14	361
CC13	BC2	3699
	BC6	312
	BC9	310
	BC12	370
	BC15	28
CC14	BC12	231
CC15	BC5	289

The date, time, temperature, light and CO_2 were considered the contextual attributes while relative humidity and humidity ratio were considered the behavioral attributes. This classification was applied because the relative humidity depends on the temperature. Also, the humidity ratio is computed using the temperature value. The dataset includes only normal tuples so artificial anomalies were injected in order to measure the accuracy of the proposed algorithm.

G. Implementation

The proposed algorithm was implemented in Python. K-Means clustering technique was implemented through the sklearn.cluster module which includes popular unsupervised clustering algorithms (Scikit-Learn, 2018a; 2018b). Number of clusters ranged from 2 to 20 as an input to the elbow and silhouette methods.

H. Results

The optimal number of clusters was 13 in case of applying the silhouette method and 15 in case of the elbow method. The resulting number of clusters is nearly equal from both methods so we decided to use 15 clusters. The results of the intersection between contextual and behavioral clusters are shown in Table 2.

In Table 2, CCi denotes the contextual clusters ranging from i = 1 to 15 while BCi denotes the behavioral clusters ranging from i = 1 to 15. The first column in Table 1 contains the contextual clusters numbered from CC1 to CC15 denoting the 15 clusters. The second column denotes the behavioral clusters with which the contextual clusters intersect with. The third column contains the number of points which result from the intersection between the corresponding contextual and behavioral clusters. This number refers to the number of data tuples in which the contextual and behavioral clusters have the same row identifier number.

A threshold of value 200 was set to detect the contextual anomalies. If the number of data points resulting from the intersection between the contextual and behavioral clusters are less than 200, the members of these clusters are considered anomalies. Anomalies usually belong to small clusters, so the members of these clusters are considered the contextual anomalies. The members of the clusters highlighted in red are the contextual anomalies.

The accuracy of detecting contextual anomalies was computed as follows:

$$\begin{aligned}&\mathit{Accuracy\ of\ detecting\ contextual\ anomalies}\\&=\frac{\mathit{number\ of\ detected\ contextual\ anomalies}}{\mathit{total\ number\ of\ contextual\ anomalies}}=78.5\%.\end{aligned}$$

The proposed algorithm gave promising results, it works in a different way compared to the current approaches. The current approaches either merge the contextual and behavioral attributes or create several contexts and apply the anomaly detection techniques to each context separately. On the other hand, the proposed algorithm grouped the similar data tuples in the same clusters so the members of the small clusters represent the contextual anomalies.

V. Conclusion

The IoT notion gained popularity throughout the few recent years. IoT is the integration of physical objects that are coupled with software, sensors and network connectivity, which allow them to capture and transfer data. The IoT paradigm faces a lot of obstacles varying from data management to security issues. A substantial challenge is the detection of contextual anomalies from sensors' data. A contextual anomaly is a data point that varies in the context it exists in. Detection of contextual anomalies in IoT is a problematic task because the context has to be taken into consideration in the anomaly detection process. In this paper, a novel clustering based contextual anomaly detection algorithm was proposed. At first, the contextual and behavioral attributes were separated then K-Means clustering technique was applied on the contextual and behavioral attributes separately. After that, the intersection between the contextual and behavioral clusters was used to discover the contextual anomalies. The proposed algorithm was tested on a real room occupation dataset and it showed promising results. The accuracy of detecting contextual anomalies was 78.5%.

VI. References

Arroyo, A., A. Herrero, V. Tricio and E. Corchado, 2017. Analysis of meteorological conditions in Spain by means of clustering techniques. J. Applied Logic, 24: 76-89. DOI: 10.1016/j.jal.2016.11.026

Berrocal, J., J. Garcia-Alonso, J.M. Murillo and C. Canal, 2017. Rich contextual information for monitoring the elderly in an early stage of cognitive impairment. J. Pervasive Mobile Comput., 34: 106-125. DOI: 10.1016/j.pmcj.2016.05.001

Candanedo, L.M. and V. Feldheim, 2016. Accurate occupancy detection of an office room from light, temperature, humidity and CO_2 measurements using statistical learning models. Energy Build., 112: 28-39. DOI: 10.1016/j.enbuild.2015.11.071

Elbouanani, S., M.A.E. Kiram and O. Achbarou, 2015. Introduction to the internet of things security: Standardization and research challenges. Proceedings of the 11th International Conference on Information Assurance and Security, Dec. 14-16, IEEE Xplore Press, Marrakech, Morocco, pp: 32-37. DOI: 10.1109/ISIAS.2015.7492741

Han, J., M. Kamber and J. Pei, 2012. Outlier Detection. In: Data Mining: Concepts and Techniques, Elsevier, Netherlands, pp: 544-548.

Hand, D., H. Mannila and P. Smyth, 2001. Principles of Data Mining. 1st Edn., The MIT Press, Cambridge, ISBN-10: 026208290X, pp: 546.

Hayes, M.A. and M.A.M. Capretz, 2014. Contextual anomaly detection in big sensor data. Proceedings of the IEEE International Congress on Big Data, Jun. 27-Jul. 2, IEEE Xplore Press, Anchorage, AK, USA, pp: 64-71. DOI: 10.1109/BigData.Congress.2014.19

Kosek, A.M., 2016. Contextual anomaly detection for cyber-physical security in Smart Grids based on an artificial neural network model. Proceedings of the Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids, Apr. 12-12, IEEE Xplore Press, Vienna, Austria, pp: 1-6. DOI: 10.1109/CPSRSG.2016.7684103

Kraijak, S. and P. Tuwanut, 2015. A survey on internet of things architecture, protocols, possible applications, security, privacy, real-world implementation and future trends. Proceedings of the IEEE 16th International conference on Communication Technology, Oct. 18-20, IEEE Xplore Press, Hangzhou, China, pp: 26-31. DOI: 10.1109/ICCT.2015.7399787

Leach, M.J.V., E.P. Sparks and N.M. Robertson, 2014. Contextual anomaly detection in crowded surveillance scenes. Patt. Recog. Lett., 44: 71-79. DOI: 10.1016/j.patrec.2013.11.018

Liu, Q., R. Klucik, C. Chen, G. Grant and D. Gallaher *et al*., 2017. Unsupervised detection of contextual anomaly in remotely sensed data. Remote Sens. Environ., 202: 75-87. DOI: 10.1016/j.rse.2017.01.034

Meng, Y., J. Liang, F. Cao and Y. He, 2018. A new distance with derivative information for functional k-means clustering algorithm. Inform. Sci., 463-464: 166-185. DOI: 10.1016/j.ins.2018.06.035

Nalbandian, S., 2015. A survey on internet of things: Applications and challenges. Proceedings of the International Congress on Technology, Communication and Knowledge, Nov. 11-12, IEEE Xplore Press, Mashhad, Iran, pp: 165-169. DOI: 10.1109/ICTCK.2015.7582664

Oh, Y. and Y. Kim, 2017. A hybrid cloud resource clustering method using analysis of application characteristics. Proceedings of the IEEE 2nd International Workshops on Foundations and Applications of Self* Systems, Sept. 18-22, IEEE Xplore Press, Tucson, AZ, USA, pp: 295-300. DOI: 10.1109/FAS-W.2017.162

Radon, A.N., K. Wang, U. Glasser, H. When and A. Westwell-Roper, 2015. Contextual verification for false alarm reduction in maritime anomaly detection. Proceedings of the IEEE International Conference on Big Data (Big Data), Oct. 29-Nov. 1, IEEE Xplore Press, Santa Clara, CA, USA, pp: 1123-1133. DOI: 10.1109/BigData.2015.7363866

Rajeswari, A.M., S.K. Yalini, R. Janani, N. Rajeswari and C. Deisy, 2018. A comparative evaluation of supervised and unsupervised methods for detecting outliers. Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies, Apr. 20-21, IEEE Xplore Press, Coimbatore, India, pp: 1068-1073. DOI: 10.1109/ICICCT.2018.8473123

Ray, S., Y. Jin and A. Raychowdhury, 2016. The changing computing paradigm with internet of things: A tutorial introduction. IEEE Design Test, 33: 76-96. DOI: 10.1109/MDAT.2016.2526612

Rose, D., 2014. Enchanted Objects: Design, Human Desire and the Internet of Things. 1st Ed., Simon and Schuster, New York, ISBN-10: 1476725632, pp: 304.

Scikit-Learn, 2018a. API Reference. Scikit-learn. https://scikit-learn.org/stable/modules/classes.html#module-sklearn.cluster

Scikit-Learn, 2018b. sklearn.cluster.Kmeans. Scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

Thah, P.H. and I.S. Sitanggang, 2016. Contextual outlier detection on hotspot data in Riau Province using k means algorithm. Proc. Environ. Sci., 33: 258-268. DOI: 10.1016/j.proenv.2016.03.077

Wang, K., X. Qi, H. Liu and J. Song, 2018. Deep belief network based k-means cluster approach for short-term wind power forecasting. Int. J. Energy, 165: 840-852. DOI: 10.1016/j.energy.2018.09.118

Yue, Z., W. Sun, P. Li, M.U. Rehman and X. Yang, 2015. Internet of things: Architecture, technology and key problems in implementation. Proceedings of the 8th International Congress on Image and Signal Processing, Oct. 14-16, IEEE Xplore Press, Shenyang, China, pp: 1298-1302. DOI: 10.1109/CISP.2015.7408082