International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 9 • Number 46 • 2016

Air Pollution Data Analysis Using Time Series Clustering for IOT

Divya Joshi[a], A Sai Sabitha[b] and Shilpi Sharma[c]

[a-c]Aset, Department of Computer Science and Engineering, Amity University Uttarpradesh, Noida, India. Email: [a]divyaj2021@gmail.com

Abstract: In today's era as the growth of technology and industries has made the life style easier but adversely affecting the environmental conditions. Air pollution is one of the major global issues which need to be resolved. Degradation of air quality in cities is the result of an increase in urbanization and industries and because of poor control over emissions. The huge volume of pollution data collected needs to be understood and analyzed. The time series pollution data can be used to extract patterns (seasonal) or novel pattern by techniques like clustering, prediction or forecasting, segmentation etc. In this research time series data is used to analyze the pollution data using time series clustering technique K-Mean and X-Mean.

Keywords: Internet of Things, Data Mining, Clustering, Time series.

1. INTRODUCTION

The technological growth leads towards the direction of Internet of Thing (IoT). The Internet of Things referred as the next generation of internet which contains trillions of nodes representing various devices/objects. These objects represent small ubiquitous devices and Handhelds to large web servers and supercomputer clusters. IoT is an integrated part of future internet technology, which defines a dynamic global network infrastructure with self configuring capabilities based on interoperable communication protocols. The IoT enables nodes to interact and communicate among themselves and with environment by exchanging data sensed about the environment. These entities act autonomously in the physical world and influencing it by running processes and create services with the minimal human intervention.

The Internet of Things consists of complex data types which includes sensor data, radio frequency data (RFID), video data and image data. The data in IoT can be categorized as: Radio frequency identification data stream, descriptive data, positional data, environmental and sensor data etc. Vast amounts of data are generated by the environmental sensors. The major challenge is to manage, analyze and mining data in its environment.

Air pollution is the major health issue affecting the urban areas of developing and the developed countries. The air pollution affects health of living beings in different aspects as there are different sources. Not only the ambient air quality affects but also the indoor air quality in both the rural and urban areas is the major concern. Air particles in the atmosphere when present above the specified level cause the pollution. These air pollutants

when inhaled affects the human health by affecting the lungs and respiratory system; also taken up by the blood and pumped all round the body. These pollutants are also deposited on the soil, plants and in the water. This is a very critical issue nowadays as it affects the population in different aspects such as health, natural disaster etc. In this research work the focus is to understand and analyze the huge volume of time series pollution data collected from Central Pollution Control Board using data mining techniques. The analysis is conducted using K-mean and X-mean clustering algorithm for time series data.

The paper is structured as follows: Theoretical background, Methodology, Experimental setup, Result analysis and conclusion and future scope.

2. THEORETICAL BACKGROUND

2.1. IoT and City Governance

The Internet of Things can be used in the field of city governance. It is used in all scenarios for public services by governments. The IoT can uniquely address problems of pollution like air pollution, water and noise, landfill waste and deforestation. The sensor devices help to monitor the environmental impact of city data like pollution data, population data, air quality data and waste disposal. IoT in city governance includes creating frameworks of best practice for smart city projects, collaboration of city services, risk management and development of a city protocol to effectively manage the urban IoT.

The major challenge is to manage and analyze this huge volume of data. By using big data analytics, researchers can predict how many residents are likely to move away from the city, and which factors of urban life lead to this decision of residents. IoT has introduced data mining techniques to manage these data in order to find an efficient pattern.

2.2. Environmental Monitoring

Environmental monitoring with Internet of Things is encouraging researchers with different revolutionary ideas. It involves planning, monitoring and conservation of natural resources. It results in an enhanced protection and resource management. It deals with the climate change, ocean and coastal management, natural resource management and biodiversity conservation. In Environmental monitoring, connected IoT sensors collect data that are emitted from factories, vehicles, detects forest fires and senses the temperature change. From the monitoring of air pollution to the real time monitoring of water quality through the IoT sensors that sends information via a GPRS network.

Following are the major area of environmental monitoring through IoT technology:-

Air quality monitoring, Water quality monitoring, Atmosphere and soil condition monitoring, Animal control, Forest fire detection, Earthquake or tsunami warning Fishing.

The research areas of Internet of Things for environmental monitoring (Refer Figure 1) are given below:

2.2.1. Pollution Monitoring

Nowadays, with the growth of industries, a steady change in the composition of atmospheric elements is found due to the combustion of fossil fuels used for the energy generation. This is due to the poor control of emissions during the process. This causes serious health problems for all citizens and animals in metropolitan cities. Today's one of the major environmental and health concern is air pollution. It has a significant risk factor for multiple health issues, including skin and eye infections, lung cancer, pneumonia, cough, and asthma.

S. No.	Paper title	Year	Authors	Technique
1	Smart Device to monitor water quality to avoid pollution in IoT environment	2015	Pandian et al.,[25]	IOT
2	Embedded System for Noise Pollution Monitoring using IoT Platform to create Smart Environment	2015	Sushma et al., [26]	IOT
3	Multi Model Air Pollution Estimation for Environment Planning using Data mining	2012	A. Vinayagam1 et al., [4]	Review
4	Smart and Secure Monitoring of Industrial Environments using IoT	2015	Puranik et al., [11]	IoT
5	Smart Environment Monitoring System by employing Wireless Sensor Networks on Vehicles For Pollution Free Smart Cities,	2015	Muhammad Saqib et al., [21]	IOT, Sensors
6	ICT Methodologies and Spatial Data Infrastructure for air quality management	2012	Francesco et al., [20]	IOT+ spatial data infrastructure.
7	Development of an IoT Environmental Monitoring Application with a Novel Middleware for resource constrained devices	2014	Salvatore Gaglio et. al. [15]	IOT using symbolic processing
8	Demonstration Abstract: Participatory Sensing enabled Environmental Monitoring in Smart	2014	Florian Zeiger & Marco F. Huber, [14]	Data analytics, IOT
9	Using User Generated Online Photos to Estimate and monitor air pollution in major cities	2015	Yuncheng Li et.al.[10]	Image analytics
10	Vehicular Pollution Monitoring Using IoT	2014	Souvik Manna[28]	IoT
11	Smart and Secure Monitoring of Industrial Environments using IoT	2015	Shruthi Puranik et al., [11]	IoT
12	Review Paper on Air Pollution Monitoring system	2015	Snehal & Priya, [13]	Real Time monitoring
13	Managing air quality by 'Data Mining'	2012	Keith McCabe et. al. [16]	Generalised Additive Models (GAMs), Bivariate polar plots

Figure 1: IoT techniques for environmental monitoring

By considering the major issues caused by the air pollution, it is very much required to monitor and control the air pollution. One of the suggested ways is to monitor the exceeding levels of air pollutants and by taking appropriate actions to control it. Several data mining techniques have been used to monitor air pollution.

2.3. IoT and Data Mining

Data mining is the process of analyzing and knowledge discovery from a massive set of data. The major objective of data mining is to find efficient patterns from the huge volume of data received from the Internet of Things devices (IoT sensors). Knowledge discovery, pattern analysis is the main tasks of data mining for it. In city governance data mining is used to discover public needs and decision making with automated systems,

improve service performance. The data mining techniques like classification, clustering and time series analysis used to solve problems in this area. For smart cities, city incident information management system can integrate data mining techniques to provide a comprehensive assessment of the impact of natural disasters on the agriculture.

2.3.1. Classification Technique

Classification techniques are used to predict a certain outcome based on a given input parameters. The classification methods like Decision tree can be used to classify whether a particular region is polluted or not. Based on the pollutant values (e.g. PM10, SO2). We can classify whether a region comes under the pollution categories like highly risky, Risky, Moderate and healthy. Data mining techniques can be used for policy making [24] to manage the pollution. The Classification and Regression Tree (CART) technique uses specialized software to identify air quality or meteorological variables which are strongly correlated with the ambient pollution levels. These parameters are then used to predict the future, pollution level based on the air quality parameters. Following are the classification techniques which are identified.

1. Naïve Bayes and Bayesian Belief Networks
2. Decision tree and KNN
3. Rule based classification
4. Genetic algorithm and SVM

2.3.2. Clustering Technique

Data mining Clustering groups data set into subsets in such a manner that similar type of data is grouped together, while different data set belongs to different groups. Clustering techniques like K- Means and Hierarchical clustering are used to determine the city cluster which are highly polluted. These techniques uses a distance function based on which the data clusters are calculated. Following are the main clustering techniques which are identified under data mining algorithms.

1. Partition Based(K-Mean, X-Mean)
2. Density Based(DBSCAN, OPTICS)
3. Hierarchical Clustering(AGNES, DIVISIVE)
4. Grid Based(STING)
5. Model Based(EM, Neural Network)

2.3.3. Association Analysis

Association techniques like Apriori algorithms are used to find association patterns between different pollutants. Following are the association techniques:

1. Pattern growth approach
2. Apriori based algorithm (Hashing, Sampling and Partitioning)
3. Vertical data format

The researches carried out under various data mining techniques are given below (Refer Figure 2).

S. No.	Research Topic	Year	Authors	Technique
1	A New Air Quality Forecasting Model Using Data Mining	2015	Min Huang et al. [1]	ANN
2	Air Pollution Monitoring & Tracking System Using Mobile Sensors and Analysis of Data Using Data Mining	2012	Umesh M et al., [3]	Association rule data mining technique, Apriori algorithm.
3	Unsupervised system to classify SO2 pollutant concentrations in Salamanca, Mexico	2012	J.M. Barrón et al.,[9]	SOM, Neural Network
4	Overview of data mining technique for WSN based air pollution detection system	2014	Snehal Sakarde et al., [22]	Hierarchical Clustering Algorithm, K- means Clustering Algorithm
5	Data Mining industrial air pollution data for trend analysis and air quality index assessment using a back end AQMS application software	2014	E.O.Ofoegbu et.al., [27]	Time series
6	Data mining to aid policy making in air pollution management	2014	Sheng-& Shue, [24]	Multi-scale analysis; Self-organization neural network;
7	An Integrated System for Regional Environmental Internet of Things	2014	Shifeng Fang & Li Da Xu, [23]	Big data Analysis
8	Indoor Air Monitoring Platform and Personal Health Reporting System	2015	Kin-Fai Ho et al., [19]	Data Mining; Date Capturing Platform
9	Indoor Air Quality Monitoring System for Smart Buildings	2014	Xuxu Chen et al., [18]	Regression, ANN
10	Air Pollution Monitoring and Mining Based system	2012	Yajie Ma et al., [12]	Distributed Data Mining
11	Smart and Secure Monitoring of Industrial Environments .\	2012	J.M. Barrón-Adame et al., [9]	SOM

Figure 2: Data mining techniques for pollution analysis

2.3.4. K-Means Clustering

K-Means clustering is the unsupervised learning algorithm that provides a solution to the clustering problem. The algorithm follows a set of steps to cluster a given dataset into a determined number of clusters. The main idea behind this technique is to group the similar data items based on their distance from the centroids. The centroids are the centre of each cluster.

2.3.5. X-Means Clustering

X-mean algorithm is an extended version of K-mean algorithms which overcomes the drawbacks of K-Means of providing an efficient estimation of number of clusters.

3. OBJECTIVE OF THE RESEARCH

The main focus of the research is on improving the quality of life in urban cities which is based on projected growth in pollution. As pollution growth in urban cities are related to power, water distribution, waste disposal, health, education and pollution. Currently, Environmental pollution and global warming is the key issue, there is a need to understand the pollution level of various urban cities.

Since a vast amount of data is collected, the major challenge is to understand and analyze this huge volume of pollution data. The objective of the research is to analyze the time series air pollution data collected using sensors during a specific time period (winter season) to find clusters of pollutant during this period using a suitable clustering technique. To provide an efficient solution to problems in data mining techniques like K-Mean, a well established and a simple technique that has been used for various applications and X-mean which scales better in performance are used.

4. CHALLENGES OF DATA MINING WITH IOT:

The various challenges of Data mining with IOT are as follows:

1. **Data Mining Algorithm:** This challenge involves selection of a suitable data mining algorithm to handle huge volume of data.
2. **Selection of Data Mining Model:** To manage and analyze massive set of data generated by an IoT device, it is necessary to select an efficient data mining model.
3. **Accessing and Data Extraction:** As an IoT device generates large volume of data, data extraction and accessing is a major challenge.
4. **Handling of Heterogeneous Data:** Heterogeneity of data is the feature of IoT as data in IoT environment gathered from different sensors and platforms. The challenge is to handle these heterogeneous data and to analyze.
5. **Data Storage:** Data storage is the major challenge in IoT as it is a complex task to store and retrieve huge volumes of data.
6. **Database Management:** Database management is a research challenge in Internet of Things environment. Selection of database management software and tools to handle big data is an important challenge of IoT. It includes database management in local node as well as in global nodes also.

5. METHODOLOGY

4.1. Data Mining System for IoT

The data mining system for IoT (Air pollution) (Refer Figure 3) can be discussed in six layers:

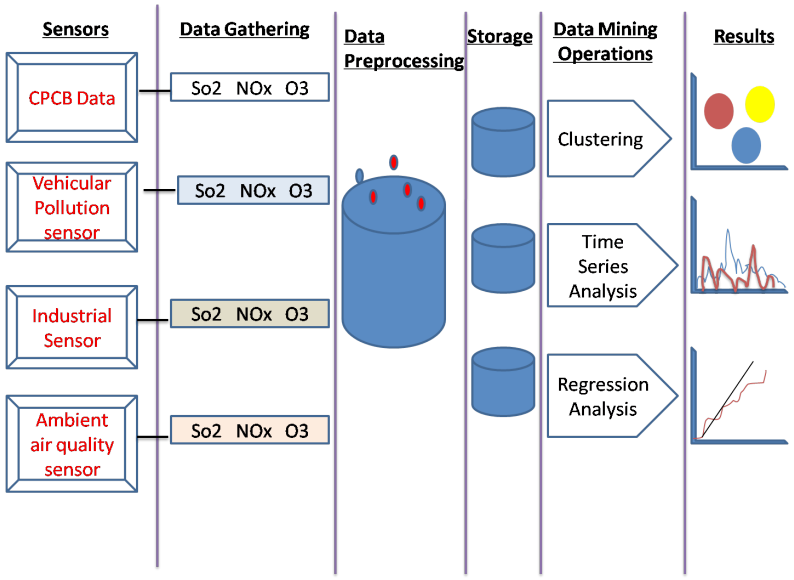


Figure 3: Data mining system for IoT in pollution

1. In the first layer IoT devices such as sensor cameras can be integrated to generate air pollution data continuously.
2. The raw data is of the structured/unstructured form. In the second layer these raw data can be gathered

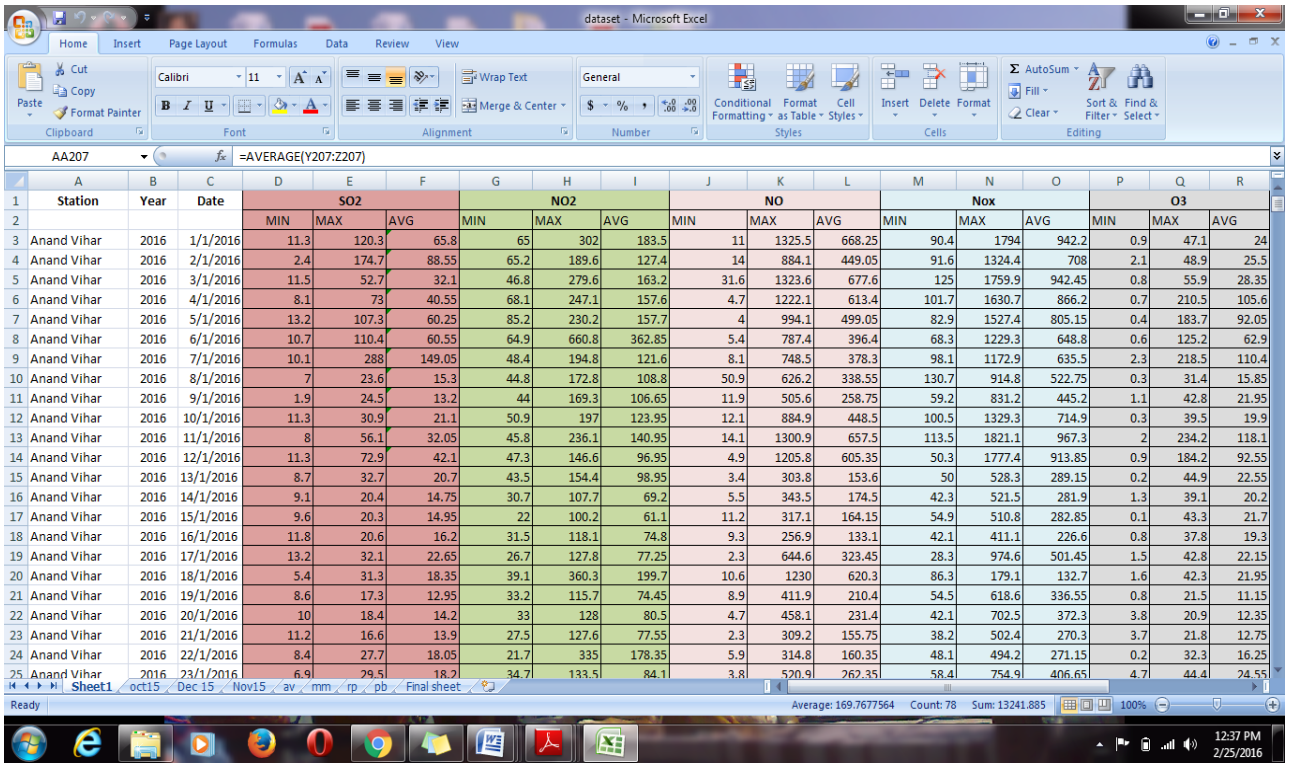
3. In the third layer the data collected can be preprocessed (cleaned, reduced, and integrated).
4. In the next layer the stored data (big data) can be processed using big data mining infrastructure.
5. Data mining functions in IoT-Air pollution data can be regression analysis, Time series, clustering.
6. The last layer shows the novel and interesting patterns discovered from huge time series data.

In this work more emphasis is given to applications of data mining technique towards city governance (pollution monitoring) and patterns are discovered in time series data.

6. EXPERIMENTAL SET UP

6.1. Data Set

The air pollution data of New Delhi are considered. The live data were collected from the website of the Central Pollution Control Board (CPCB), India [30]. The time series data were collected for the four stations of New Delhi for the analysis. The data are considered from December 2015 to February 2016 for the four different stations; Anand Vihar, Punjabi Bagh, R K Puram and Mandir Marg understand the pollution during winter season. The attributes of pollution data are NO2 (Nitrogen dioxide), SO2 (Sulpher dioxide), NO (Nitric oxide) NOx (Oxides of nitrogen), O3 (Ozone), NH3 (Ammonia), PM10 (Particulate matter 10), PM2.5 (Particulate matter 2.5). Screenshot of the data collected is given below (Refer Figure 4).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Station	Year	Date	SO2			NO2			NO			Nox			O3		
2				MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG
3	Anand Vihar	2016	1/1/2016	11.3	120.3	65.8	65	302	183.5	11	1325.5	668.25	90.4	1794	942.2	0.9	47.1	24
4	Anand Vihar	2016	2/1/2016	2.4	174.7	88.55	65.2	189.6	127.4	14	884.1	449.05	91.6	1324.4	708	2.1	48.9	25.5
5	Anand Vihar	2016	3/1/2016	11.5	52.7	32.1	46.8	279.6	163.2	31.6	1323.6	677.6	125	1759.9	942.45	0.8	55.9	28.35
6	Anand Vihar	2016	4/1/2016	8.1	73	40.55	68.1	247.1	157.6	4.7	1222.1	613.4	101.7	1630.7	866.2	0.7	210.5	105.6
7	Anand Vihar	2016	5/1/2016	13.2	107.3	60.25	85.2	230.2	157.7	4	994.1	499.05	82.9	1527.4	805.15	0.4	183.7	92.05
8	Anand Vihar	2016	6/1/2016	10.7	110.4	60.55	64.9	660.8	362.85	5.4	787.4	396.4	68.3	1229.3	648.8	0.6	125.2	62.9
9	Anand Vihar	2016	7/1/2016	10.1	288	149.05	48.4	194.8	121.6	8.1	748.5	378.3	98.1	1172.9	635.5	2.3	218.5	110.4
10	Anand Vihar	2016	8/1/2016	7	23.6	15.3	44.8	172.8	108.8	50.9	626.2	338.55	130.7	914.8	522.75	0.3	31.4	15.85
11	Anand Vihar	2016	9/1/2016	1.9	24.5	13.2	44	169.3	106.65	11.9	505.6	258.75	59.2	831.2	445.2	1.1	42.8	21.95
12	Anand Vihar	2016	10/1/2016	11.3	30.9	21.1	50.9	197	123.95	12.1	884.9	448.5	100.5	1329.3	714.9	0.3	39.5	19.9
13	Anand Vihar	2016	11/1/2016	8	56.1	32.05	45.8	236.1	140.95	14.1	1300.9	657.5	113.5	1821.1	967.3	2	234.2	118.1
14	Anand Vihar	2016	12/1/2016	11.3	72.9	42.1	47.3	146.6	96.95	4.9	1205.8	605.35	50.3	1777.4	913.85	0.9	184.2	92.55
15	Anand Vihar	2016	13/1/2016	8.7	32.7	20.7	43.5	154.4	98.95	3.4	303.8	153.6	50	528.3	289.15	0.2	44.9	22.55
16	Anand Vihar	2016	14/1/2016	9.1	20.4	14.75	30.7	107.7	69.2	5.5	343.5	174.5	42.3	521.5	281.9	1.3	39.1	20.2
17	Anand Vihar	2016	15/1/2016	9.6	20.3	14.95	22	100.2	61.1	11.2	317.1	164.15	54.9	510.8	282.85	0.1	43.3	21.7
18	Anand Vihar	2016	16/1/2016	11.8	20.6	16.2	31.5	118.1	74.8	9.3	256.9	133.1	42.1	411.1	226.6	0.8	37.8	19.3
19	Anand Vihar	2016	17/1/2016	13.2	32.1	22.65	26.7	127.8	77.25	2.3	644.6	323.45	28.3	974.6	501.45	1.5	42.8	22.15
20	Anand Vihar	2016	18/1/2016	5.4	31.3	18.35	39.1	360.3	199.7	10.6	1230	620.3	86.3	179.1	132.7	1.6	42.3	21.95
21	Anand Vihar	2016	19/1/2016	8.6	17.3	12.95	33.2	115.7	74.45	8.9	411.9	210.4	54.5	618.6	336.55	0.8	21.5	11.15
22	Anand Vihar	2016	20/1/2016	10	18.4	14.2	33	128	80.5	4.7	458.1	231.4	42.1	702.5	372.3	3.8	20.9	12.35
23	Anand Vihar	2016	21/1/2016	11.2	16.6	13.9	27.5	127.6	77.55	2.3	309.2	155.75	38.2	502.4	270.3	3.7	21.8	12.75
24	Anand Vihar	2016	22/1/2016	8.4	27.7	18.05	21.7	335	178.35	5.9	314.8	160.35	48.1	494.2	271.15	0.2	32.3	16.25
25	Anand Vihar	2016	23/1/2016	6.9	29.5	18.2	34.7	133.5	84.1	3.8	520.9	262.35	58.4	754.9	406.65	4.7	44.4	24.55

Figure 4: Data set for the analysis

6.2. Data Analysis Model

For the analysis data mining tool rapid miner was considered. The time series data were loaded into clustering tool. The screenshot of the clustering operators is shown below (Refer Figure 5).

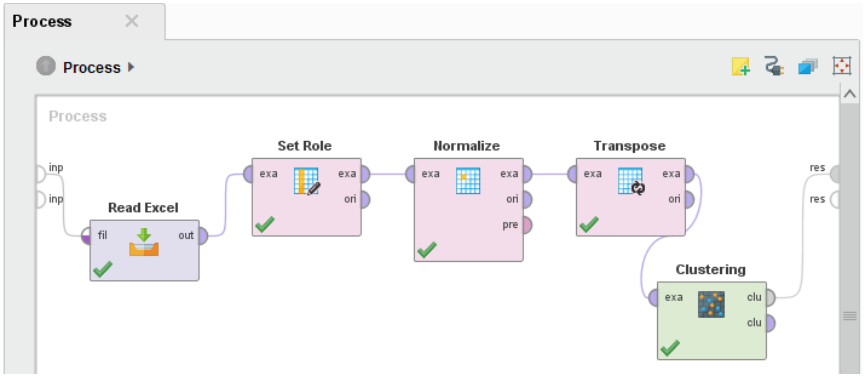


Figure 5: Clustering tool in Rapid Miner

From the above model the clusters were identified. The process was repeated using clustering operator for X-Mean. The results of the cluster model are shown in Figure 6 given below:

Cluster Model –X Mean
Cluster 0: 1 items
Cluster 1: 1 items
Cluster 2: 6 items
Total number of items: 8

Result History | ExampleSet (X-Means) | Cluster Model (X-Means)

ExampleSet (8 examples, 2 special attributes, 237 regular attributes) Filter (8 / 8 examples): all

Data | Statistics | Charts

Row No.	id	cluster	att_500.0	att_501.0	att_502.0	att_503.0	att_504.0	att
1	SO2	cluster_0	0.536	1.071	-0.256	-0.057	0.406	0.4
2	NO2	cluster_2	1.732	0.427	1.260	1.130	1.132	5.9
3	NO	cluster_2	2.491	1.188	2.547	2.165	1.485	0.8
4	Nox	cluster_2	2.329	1.310	2.330	1.998	1.732	1.0
5	O3	cluster_1	-1.113	-1.081	-1.020	0.626	0.337	-0.2
6	NH3	cluster_2	0.794	0.315	0.157	0.814	0.942	1.1
7	PM10	cluster_2	1.355	1.241	1.098	1.087	2.355	0.7
8	PM2.5	cluster_2	1.070	0.879	0.582	2.123	1.790	2.2

Cluster Model- K Mean
Cluster 0: 2 items
Cluster 1: 6 items
Total number of items: 8

Result History | ExampleSet (Clustering) | Cluster Model (Clustering)

ExampleSet (8 examples, 2 special attributes, 237 regular attributes) Filter (8 / 8 examples): all

Data | Statistics | Charts

Row No.	id	cluster	att_500.0	att_501.0	att_502.0	att_503.0	att_504.0	att
1	SO2	cluster_0	0.536	1.071	-0.256	-0.057	0.406	0.4
2	NO2	cluster_1	1.732	0.427	1.260	1.130	1.132	5.9
3	NO	cluster_1	2.491	1.188	2.547	2.165	1.485	0.8
4	Nox	cluster_1	2.329	1.310	2.330	1.998	1.732	1.0
5	O3	cluster_0	-1.113	-1.081	-1.020	0.626	0.337	-0.2
6	NH3	cluster_1	0.794	0.315	0.157	0.814	0.942	1.1
7	PM10	cluster_1	1.355	1.241	1.098	1.087	2.355	0.7
8	PM2.5	cluster_1	1.070	0.879	0.582	2.123	1.790	2.2

Figure 6: Cluster Model of X-Means & M-Mean

In the first experiment clustering using X-Mean 3clusters is formed. In K-Mean two clusters were forms.

7. ANALYSIS

Case study 1: X-Mean

The time series data loaded was standardized by performing Z- transformation of the values, so that mean is zero and standard deviation is 1. The data is further transposed and the clustering analysis was done on the transposed data. The cluster _0 had SO2 pollutant. O3 was found in cluster_1. The remaining pollutants were found in cluster_2 as shown in the table given above (Refer Figure 7(a)).

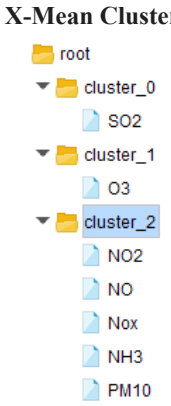


Figure 7: (a) Cluster set of X-Mean

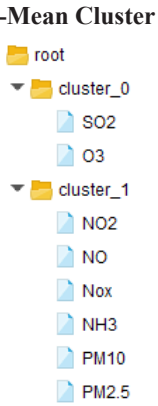


Figure 7: (b) Cluster set of K-mean

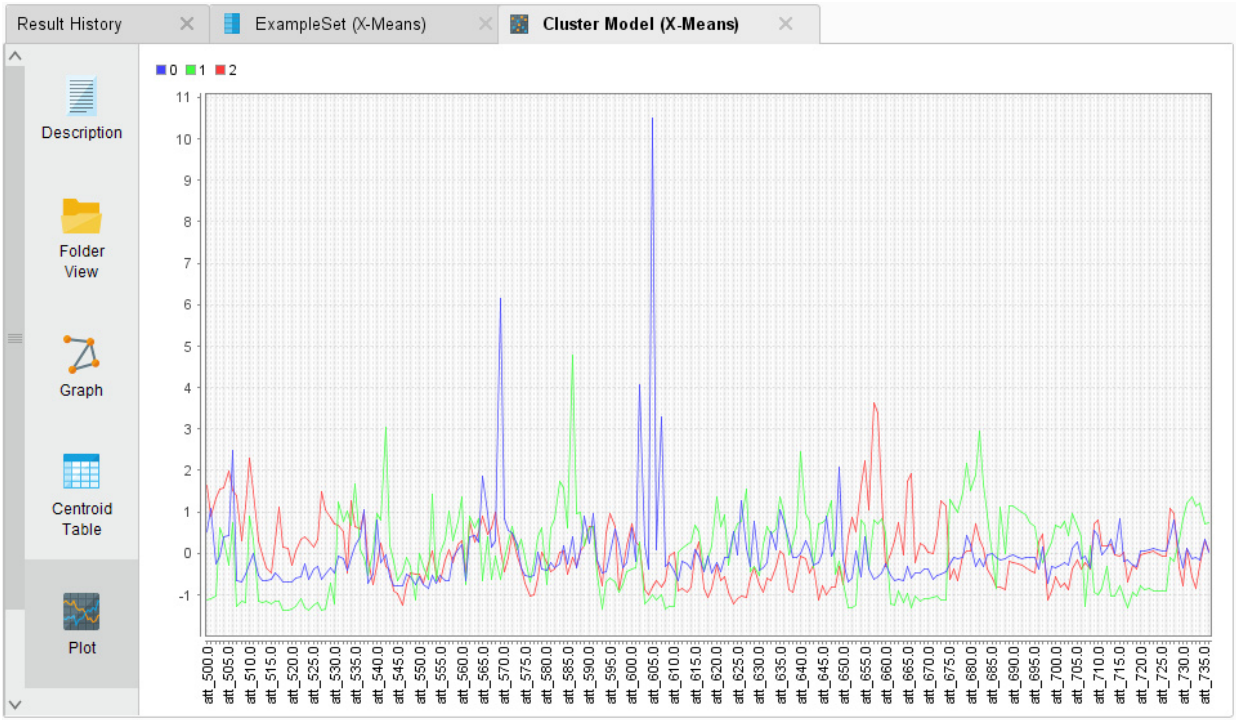


Fig. 8 Parallel plot of X-mean analysis

According to the parallel plot (Refer Fig. 8) of the cluster time series data it was found during the winter season the pollutant SO2 of cluster_0 (represented as blue line) was very high (att_605 to att_610) during Jan 13,2016 to Jan 18, 2016. The reason could be due to peak winter time and festival time of north India (Festivals like Lohri etc.). The Cluster_2 (NO2, NO, NOx, NH_3) were found to be constant except a little rise in end of Jan 2016.

Case study 2: K-Mean

The time series data is processed through the data mining model in rapid miner for K-Mean. The data processed through the transposition and normalization to the K-mean clustering operator in the tool. The data is clustered into two clusters after the processing. Cluster_0 contains two pollutants SO2 and O3. The remaining pollutants were found in Cluster_1 as shown in the above table (Refer Fig. 7(b)).

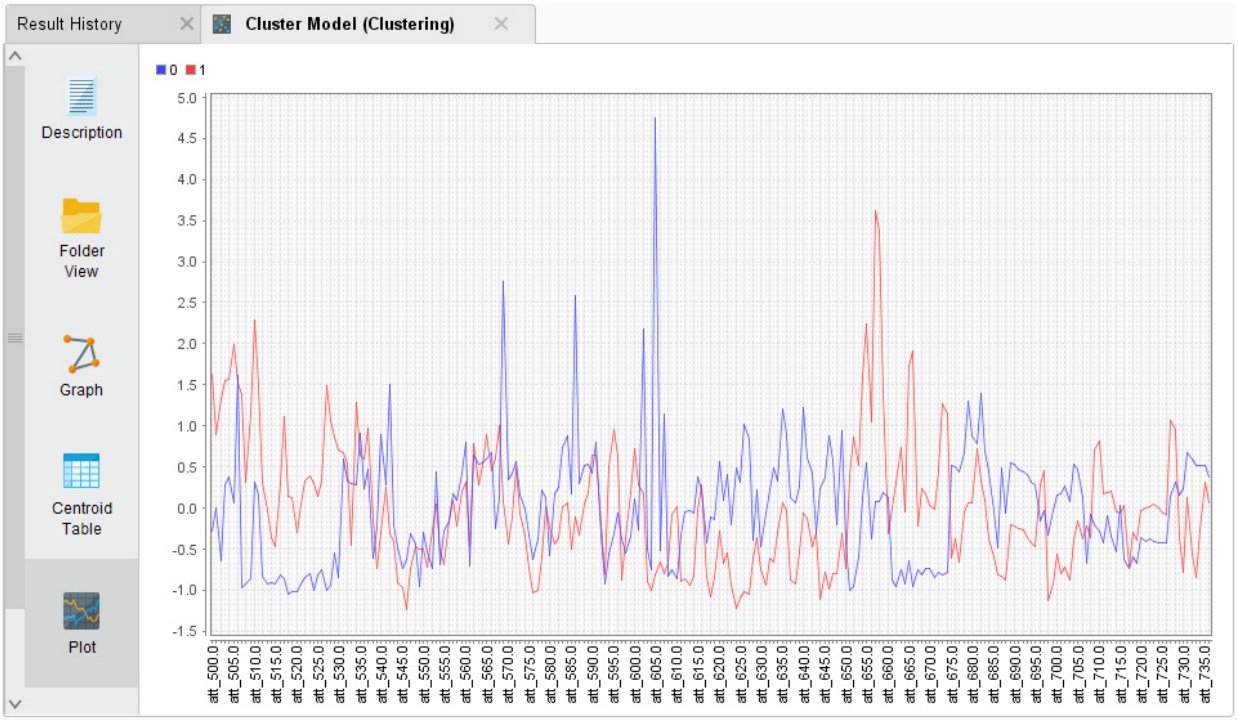


Figure 9: Parallel plot of K-mean analysis

According to the parallel plot (Refer Figure 9) of K-mean analysis the pollutant SO2 and O3 of Cluster_0 were found high during the Jan 4, 2016 to Jan 9, 2016. From Jan 13, 2016 to Jan 18, 2016 there is sudden rise of pollutants SO2 and O3 due to festival season.

During the Feb 12, 2016 to Feb 17, 2016 a rise was found in the pollution level of pollutant of cluster_1.

8. CONCLUSION AND FUTURE SCOPE

In this work time series clustering analysis was used to study the patterns of pollution data of urban city. Data mining techniques served as a good standard to discover the pollutant level during a time interval. X-mean algorithm was found to be better than K-mean as the number of clusters efficiently partitioned. The X-mean algorithm scaled better than the K-Means since it uses heuristic approach to determine the number of centroids. It uses BIC (Bayesian Information Criteria) for a better partition. Time series analysis can further be used to predict/forecast for the next few days/months. Classification techniques like regression, neural network, and support vector machine can be used as a modeling for prediction.

Acknowledgement

The pollution dataset has been collected from Central Pollution Control Board (CPCB) web site [30]. We acknowledge our sincere thanks to CPCB.

REFERENCES

[1] Huang M, Zhang T, Wang J, Zhu L. A new air quality forecasting model using data mining and artificial neural network. InSoftware Engineering and Service Science (ICSESS), 2015 6th IEEE International Conference on 2015 Sep23(pp.259-262).

[2] Dlodlo, Nomusa. "Adopting the internet of things technologies in environmental management in South Africa." (2012).

[3] Lanjewar UM, Shah JJ. Air pollution monitoring & tracking system using mobile sensors and analysis of data using data mining. International Journal of Advanced Computer Research. 2012;2(4):19-23.

[4] Vinayagam A, Kavitha C, Thangadurai K. Multi Model Air Pollution Estimation for Environmental Planning Using Data Mining.

[5] Fang S, Da Xu L, Zhu Y, Ahati J, Pei H, Yan J, Liu Z. An integrated system for regional environmental monitoring and management based on internet of things. Industrial Informatics, IEEE Transactions on. 2014 May;10(2):1596-605..

[6] Boulos MN, Al-Shorbaji NM. On the Internet of Things, smart cities and the WHO Healthy Cities. International journal of health geographics. 2014 Mar 27;13(1):1.

[7] Zhao J, Zheng X, Dong R, Shao G. The planning, construction, and management toward sustainable cities in China needs the Environmental Internet of Things. International Journal of Sustainable Development & World Ecology. 2013 Jun 1;20(3):195-8..

[8] Sakarde S, Chaudhary MM, Gode MS. Overview of data mining technique for WSN based air pollution detection system. InInternational Journal of Engineering Development and Research 2014 Mar (Vol. 2, No. 1 (March 2014)). IJEDR.

[9] Barrón-Adame JM, Cortina-Januchs MG, Vega-Corona A, Andina D. Unsupervised system to classify SO 2 pollutant concentrations in Salamanca, Mexico. Expert Systems with Applications. 2012 Jan 31;39(1):107-16.

[10] Li Y, Huang J, Luo J. Using user generated online photos to estimate and monitor air pollution in major cities. InProceedings of the 7th International Conference on Internet Multimedia Computing and Service 2015 Aug 19 (p. 79). ACM.

[11] Puranik S, Mohan J, Chandrasekaran K. Smart and Secure Monitoring of Industrial Environments using IoT. InProceedings of the Third International Symposium on Women in Computing and Informatics 2015 Aug 10 (pp. 644-649). ACM..

[12] Ma Y, Richards M, Ghanem M, Guo Y, Hassard J. Air pollution monitoring and mining based on sensor grid in London. Sensors. 2008 Jun 1;8(6):3601-23.

[13] Sirsikar, Snehal, and Priya Karemore. "Review Paper on Air Pollution Monitoring system." *International Journal of Advanced Research in Computer and Communication Engineering* (2015) 4.1.

[14] Zeiger F, Huber M. Demonstration abstract: participatory sensing enabled environmental monitoring in smart cities. InProceedings of the 13th international symposium on Information processing in sensor networks 2014 Apr 15 (pp. 337-338). IEEE Press.

[15] Gaglio S, Re GL, Martorella G, Peri D, Vassallo SD. Development of an IoT Environmental Monitoring Application with a Novel Middleware for Resource Constrained Devices. InProceedings of the 2nd Conference on Mobile and Information Technologies in Medicine (MobileMed 2014).

[16] Bywaters A, White J, McCabe K, Taylor PJ, Carslaw D. Managing air quality by 'Data Mining'.

[17] Gubbi J, Buyya R, Marusic S, Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems. 2013 Sep 30;29(7):1645-60.

[18] Chen X, Zheng Y, Chen Y, Jin Q, Sun W, Chang E, Ma WY. Indoor air quality monitoring system for smart buildings. InProceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing 2014 Sep 13 (pp. 471-475). ACM.

[19] Ho KF, Hirai HW, Kuo YH, Meng HM, Tsoi KK. Indoor Air Monitoring Platform and Personal Health Reporting System: Big Data Analytics for Public Health Research. InBig Data (BigData Congress), 2015 IEEE International Congress on 2015 Jun 27 (pp. 309-312). IEEE.

[20] D'Amore F, Cinnirella S, Pirrone N. ICT methodologies and spatial data infrastructure for air quality information management. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of. 2012 Dec;5(6):1761-71.

[21] Jamil MS, Jamil MA, Mazhar A, Ikram A, Ahmed A, Munawar U. Smart Environment Monitoring System by employing Wireless Sensor Networks on Vehicles For Pollution Free Smart Cities. Procedia Engineering. 2015 Dec 31;107:480-4.

[22] Sakarde S, Chaudhary MM, Gode MS. Overview of data mining technique for WSN based air pollution detection system. InInternational Journal of Engineering Development and Research 2014 Mar (Vol. 2, No. 1 (March 2014)). IJEDR.

[23] Fang S, Da Xu L, Zhu Y, Ahati J, Pei H, Yan J, Liu Z. An integrated system for regional environmental monitoring and management based on internet of things. Industrial Informatics, IEEE Transactions on. 2014 May;10(2):1596-605.

[24] Li ST, Shue LY. Data mining to aid policy making in air pollution management. Expert Systems with Applications. 2004 Oct 31;27(3):331-40.

[25] Pandian DR, Mala K. Smart Device to monitor water quality to avoid pollution in IoT environment.

[26] Sushma Maithare and Dr.Vijaya Kumar B P. " Embedded System for Noise Pollution Monitoring using IoT Platform to create Smart Environment".

[27] Ofoegbu EO, Fayemiwo MA, Omisore MO. DATA MINING INDUSTRIAL AIR POLLUTION DATA FOR TREND ANALYSIS AND AIR QUALITY INDEX ASSESSMENT USING A NOVEL BACK-END AQMS APPLICATION SOFTWARE. International Journal of Innovation and Scientific Research. 2014 Nov 2;11(2):237-47.

[28] Manna S, Bhunia SS, Mukherjee N. Vehicular pollution monitoring using IoT. InRecent Advances and Innovations in Engineering (ICRAIE), 2014 2014 May 9 (pp. 1-5). IEEE.

[29] Han J, Kamber M, Pei J. Data mining: concepts and techniques. Elsevier; 2011 Jun 9.

[30] http://www.cpcb.gov.in/CAAQM/Auth/frmViewReportNew.aspx