International Journal of Advanced Scientific Inovation Volume 01 Issue 01, December 2020
ISSN: 2582-8436 doi:10.5281/zenodo.4641703

Real Time Environmental Time Series Data Analysis Using Influx DB

Gangadhar S [#1], Sowmya[*2]
[#] *Software Developer, TCS Research and Innovation Labs, Bangalore, India*
[#] *Software Developer, Infosys, Bangalore, India*
gangadhar.smg24@gmail.com

Abstract - Today monitoring environmental parameters have gained more importance due to the increasing security and regulatory compliance needs. So the measurement and such parameters become critically important. The main aim of this paper is to monitoring the environmental temperature and humidity. There will be some fluctuations in the environment, the analysis and visualization of such changes is difficult. The proposed system is designed using Internet of Things based embedded system(Arduino) and time series database InfluxDB. We are using DHT11 sensor to collect the data from the environment and then stores the collected data in Influxdb. For any mischievous or any anomalous situation, alert is send to the concern people. Dynamic alert is generated with use of kapacitor and also grafana is used for visualizing the results. The collected data is also exported for smart data analysis using deep neural network technique to avoid the manual operations. The experimental results shows that this system can be used effectively in smart city different applications like Smart Street light management and smart electric grids.

Keywords — Internet of Things (IoT); Time series data; Influx DB; Kapacitor, ARIMA

I. INTRODUCTION

Environmental monitoring is a process that describes the activities that need to be processed to characterize and monitor the environment quality. Environmental monitoring is used in the preparation of assessment for environmental impact. A time series data is a collection of information points are considered in timely order [1]. Time series is a sequence taken at successive time interval in particular space. Time series data analysis contains techniques to analyze time series data in order to fetch the important statistics and other features of the time series data. A time series database is a database that is minimized to handle the time series data, collection of numbers that are indexed by time.

Some of the problems in Environment monitoring is listed below.

i. There is lot of complexity in maintaining environmental statistics, measured quantities (temperature & humidity) do not follow any rules, they usually change over time and space for this reason time series analysis is very much needed.

ii. Environmental statistics has huge amount of data with in it, although there are many techniques to store and analyze these data, fast retrieval and processing is difficult using existing methods.

iii. As the data is huge, might loss the data, data retention is very difficult using existing methods and as we are analyzing time series data, frequent queries on data is very important. Existing method has lengthy process for storing and querying data.

The main aim of our paper is to build a cost effective and high performance device to analyze the time series data using arduino which monitor the environment, to build the system which store, retrieve, process the large statistical data very fast, and to build the system which achieves data retention. This paper also aims to visualize the environmental monitoring data and send notification to the users through mails.

II. RELATED WORK

A. Time series data and IOT

A time series data is a collection of information points are considered in timely order. Time series is a sequence taken at successive time interval in particular space. Time series data analysis contains techniques to analyze time series data in order to fetch the important statistics and other features of the time series data. A time series database is a database that is minimized to handle the time series data, collection of numbers that are indexed by time [2][3]. The Internet of Things (IOT) is a system of inter-networking of operating machines, vehicle building and Digital machines, objects, animals or people that are given with unique identifiers and has the capability of transferring information over a network without any requirement of human-to-human or human-to-machine interaction. A object, in the IOT, are like human heart monitor implant, it is a type of animal with a biochip, at the time when tire pressure is less an automobile is built with the sensing device to inform to the driver or any other natural or man-made things that can be allocated to an IP address and provided with the capable of transferring information over a network.

B. Literature work

Author Mohannad Ibrahim et al [4], developed a IoT designed on Smart Environmental monitoring using the

Raspberry-Pi computer. It provides monitoring services for remote areas and for ad hoc applications and earthquake detection capability can help saving millions of lives.

Author M.B. Waghmare et al [5], proposed Temperature and Humidity Analysis using Data Logger of Data Acquisition System. In this paper analyzed results of outputs are compared with the actual values of humidity and temperature by that accuracy can be determined.

Author Vijay S. Kale et al [6], states about Real Time Remote Temperature & Humidity Monitoring Using Arduino and Xbee S2. In this, Real-time data is received using Xbee S2 and it was retained and visualized to attain temperature and humidity monitoring.

Author Gurudeo Anand Tularam et al [7], briefed Rainfall and Temperature Interactions in Coastal Catchments using Time Series Analysis. In this, the autocorrelation in the data appears to have caused the low and ARIMA techniques gives time series models for autocorrelation and seasonality[8].

III. PROPOSED SYSTEM

The existing weather monitoring system for reading weather report, the systems are designed to have the great ability to read and log the temperature humidity that data is stored in MS-XL or in text or in MySQL, based on the stored data the weather can be predicted. The drawbacks of the existing system are

• Existing system does not support high write loads, large data storage and do not have a property of conserving space by deleting unwanted data.

• We cannot retrieve the data fast from the existing system. It is very difficult to visualize the real time data in the existing system.

• Existing system uses relational database that is dependent on the external system To overcome the drawbacks of the existing time series data analysis system we are proposing a new time series data analysis using sensors and time series database Influx DB.

• The Implementation of Time Series Data Analysis using arduino UNO took advantages of the storing and retrieval of time series data that is coming from temperature sensor very fast.

• The proposed system gives a best method for visualizing the time series data using grafana a visualizing metrics.

• We can also send notification like Email, slack, pager duty using grafana alerting.

The Fig.1 below shows the symbolic diagram of a system which contains the main components and functional unit of the system. To analyze the time series data Influx DB and grafana is used.

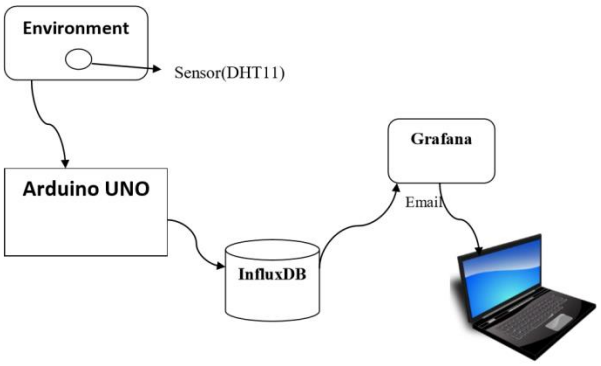

Fig 2: Symbolic diagram of a system

Influx DB is developed by Influx data, it is an open-source time series database written in go. Influx DB is designed for large availability of storage space and fast retrieval of a sequence of data in the areas like operations monitoring, internet of things sensors data, application metrics and also in real time analytics.

Grafana is an open source tool mainly designed to analyze the time series data .it provides many way to create, share, explore and to understand the data in a graphical representation it support many databases such as Influx DB, graphite, OpenTS DB etc [9].

The below Fig. 2. shows the system architecture of IOT based time series data analysis which compromises of envi ronment where the temperature and humidity to be checked to get the current temperature we are using DHT11 sensor which produce sensor data, Arduino UNO board, Influx DB database is used to save the data generating on the DHT11 sensor. The data is send to grafana to visualize and later notification is created using alerting. Arduino board with the SD card is connected to computer is booted. Arduino IDE is downloaded and sketch (programs) are written to retrieve the temperature readings from the DHT11 sensor. An API is send to the data to Influx DB which stores the data collected from the Arduino board. We query those data in Influx DB like MySQL query. When the data is moved from Influx DB to grafana we can visualize the time series data on the grafana and also we can send the alerts using alerting component in the grafana, the alerts may be Email or message etc.

Proposed system can be used to monitor the data centers for air temperature and humidity. It can be used to monitor air quality, carbon dioxide, and smoke like gasses an indoor ozone layer levels. The proposed system can be used to measure the pollutants in the city environments. This paper main scope is to focus on the time series analysis of environmental measurements for temperature and humidity. The system explains the need of environment monitoring to know the climatic changes which affects the natural phenomena.

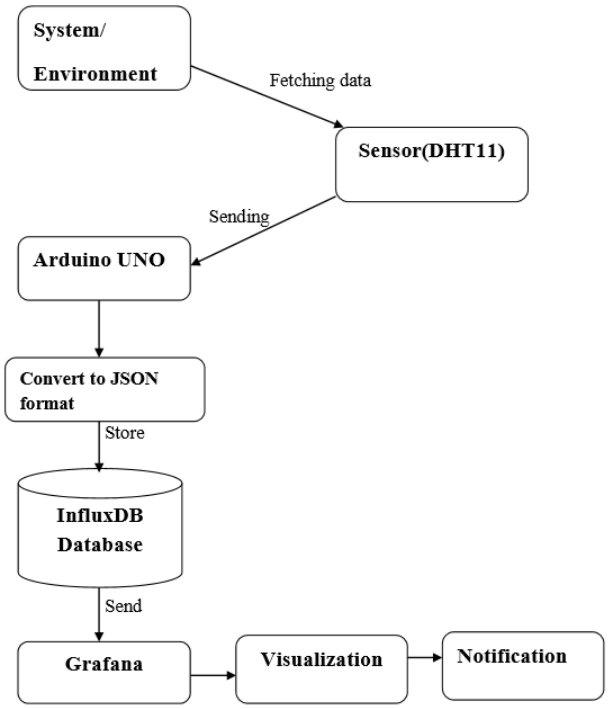

Fig 2: System Architecture

IV. IMPLEMENTATION

This process is divided into 5 modules.

1. Data collection using DHT11 sensor.
2. Fetch data from Arduino board
3. Converting data to JSON format and store it in Influx DB.
4. Visualizing using grafana.
5. Sending notification to user.

Initially, access the data from the DHT11 sensor through the Arduino board interface and connection made as shown in Fig. 3 and 4. Then the collected data will be stored on the Arduino board. The sensor data should be fetched to store it in Influx DB Fig 6. For this we need to run the code in Arduino IDE.

To store the data in influx DB, the data need to be in JSON format. The CSV data is converted to JSON using python code. Stream the data coming from arduino board, this data will be stored in the Influx DB at a regular time interval and it will be analyzed in a real time and be visualized using grafana dashboard.

After visualizing data in grafana we can send the notifications to the user, when alert changes its state it will send the notification, every alert rule support many notifications. In order to add alert rule to your notification you first need to install and configure notification channel. This can be done in the notification channel page in grafana

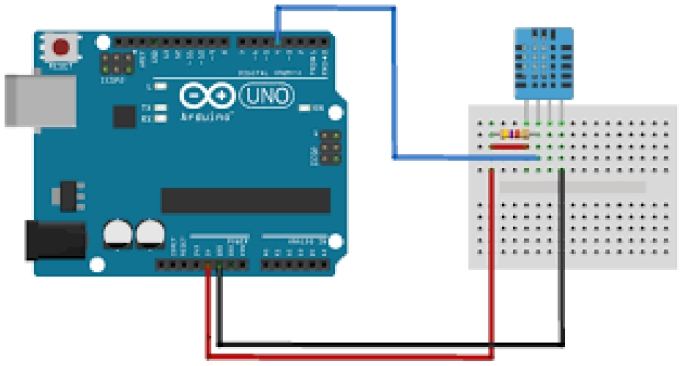

Fig. 3: Connection between Arduino and DHT11 sensor

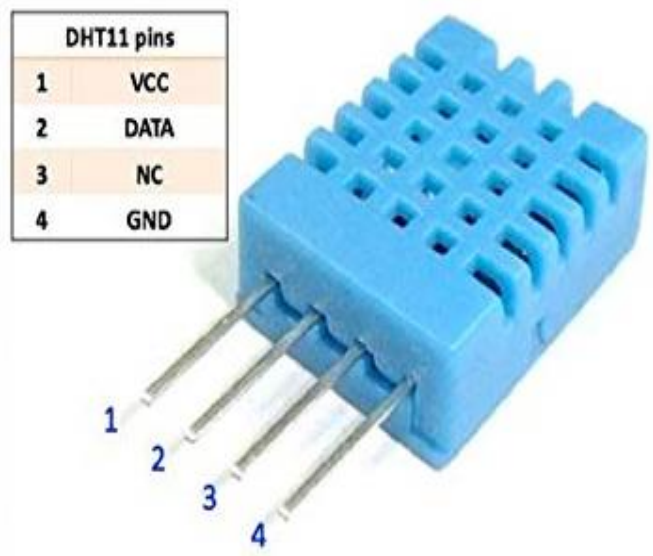

Fig. 4: DHT11 sensor

Description	Expected result	Status
Collecting data from	Gathering data	pass

DHT11 sensor		
Converting data to JSON format Data is converting into JSON to store data in Influx DB	Data is converting	pass
Store data to Influx DB Data is storing in influx DB	Data is storing	pass
Visualize the data	Graphing, visualization	pass
Generating the alerts Alerts are generating through mails	Alerts are generating	pass

TBALE:1

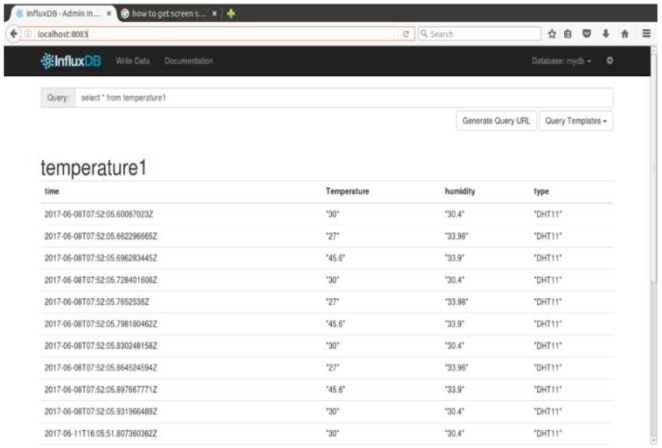

Fig. 6: Illustrating Influx DB data

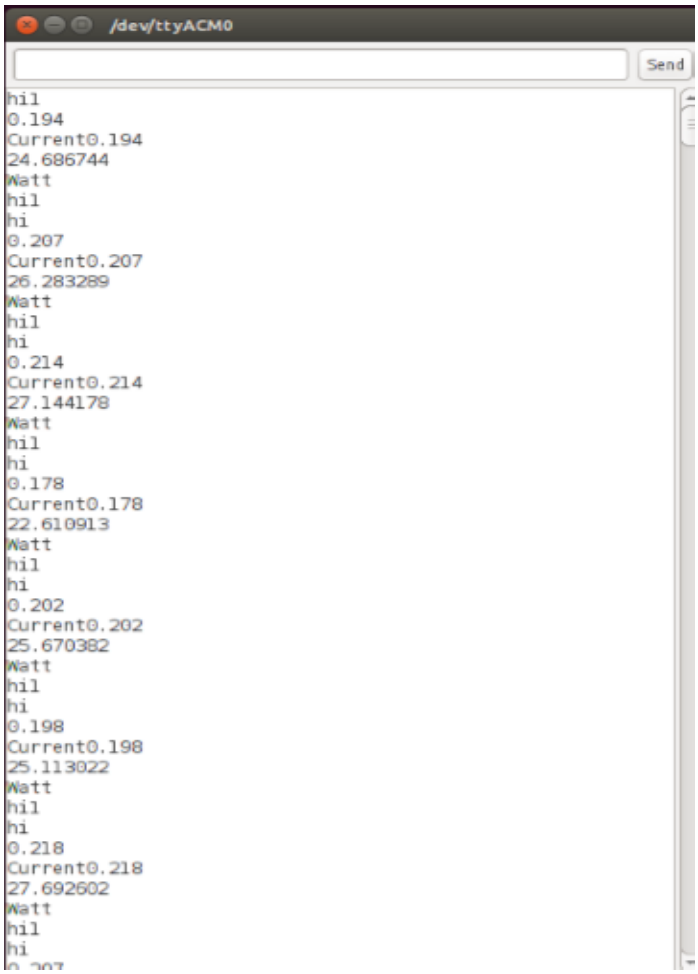

Fig. 5: Illustrating Serial monitor

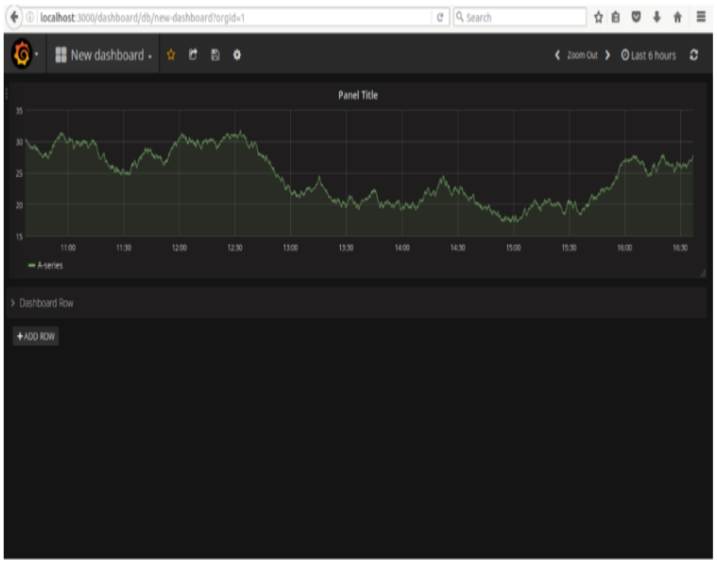

Fig. 7: Illustrating graphical representation of data

III. EXPERIMENTAL RESULTS

The proposed system need to be tested on the values that are fetched from the current sensor. This is because we are in need of those data to analyze the time series data. As we are using the DHT11 sensor it is very sensitive to the small variations in temperature and humidity, these small changes should also must be stored in database. After this is done the user needs to login to grafana and try to visualize data stored in Influx DB through graphs. Later by creating channels user can able to send notification in case of emergency. To get the data, visualize and send notifications these test cases are done.

A. Benefits of proposed system

- Very efficient way to analyze the time series data analysis.
- As we are using Influx DB to store the data getting from the sensor we can easily do query like normal SQL.
- Influx DB has the property called retention policy you can automatically handle your data retention periods.
- Accurate values of temperature and humidity can be obtained from the proposed system.
- As our system using Influx DB data base we can store & retrieve the data very fast.
- Real time analysis of data is possible from the proposed system.
- User can send the notification through Email or through other notification types in case if there is large variations in temperature and humidity.

B. Features of proposed system

- Real time data analysis.
- Data retention policy can be obtained.
- Visualize data through graphs.
- Better and faster retrieval of data Recommended font sizes are shown in Table 1.

VI. CONCLUSIONS

Traditional relational databases are suitable for most kind of applications, but not effective in terms of cost, maintenance and reliability. Influx DB provided a simple way to accomplish the required task with its features capable to serve both small scale and enterprise level applications. With the help of grafana, user can analyze the real time data through graphical representation. This process sends notification through mails, in case if there is a large variation in temperature and humidity. In this paper we made an attempt to integrate IoT with Data analysis techniques using deep neural network technique to avoid the manual operations.

REFERENCES

[1] Ramya M, Chetan Balaji, and Girish L "Environment Change Prediction to Adapt Climate-Smart Agriculture Using Big Data Analytics." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 (2015).

[2] Girish, L; Deepthi T, K.i-Manager's Journal on omputer Science; Nagercoil Vol. 6, Iss. 2, (Jun/Aug 2018): 1.DOI:10.26634/jcom.6.2.14870

[3] Vishwas Rudramurthy, Bharath T R, Gangadhar S "Applying Cloud Computing and Wireless Sensor Technology for making Smart Agriculture" International Journal for Research in Engineering Application and Management (IJREAM), page 38-42.

[4] Bartosz Balis, Marian Bubak, Daniel Harezlak, Piotr Nowakowski, Maciej Pawlik, Bartosz Wilk "Towards an Operational database for realtime environmental monitoring and early warning systems", International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland.

[5] Girish L, Gangadhar S, Bharath T R, Balaji K S, Abhishek K T "Crop Yield and Rainfall Prediction in Tumakuru District using Machine Learning" International Journal for Research in Engineering Application and Management (IJREAM), page 61-65.

[6] Nayana Y, Mr.JustinGopinath, Girish.L"DDoS Mitigation using Software Defined Network", International Journal of Engineering Trends and Technology (IJETT), V24(5),258-264 June 2015. ISSN:2231-5381. www.ijettjournal.org. published by seventh sense research group

[7] Sahana, D. S., and L. Girish. "Automatic Drug Reaction Detection Using Sentimental Analysis." , International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 5, May 2015

[8] Rashmi, T. V., and Keshava Prasanna. "Load Balancing As A Service In Openstack-Liberty." International Journal of Scientific & Technology Research 4.8 (2015): 70-73.

[9] Khanum, Salma, and L. Girish. "Meta Heuristic Approach for Task Scheduling In Cloud Datacenter for Optimum Performance." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4

[10] L, G. (2019). Anomaly Detection in NFV Using Tree-Based Unsupervised Learning Method. International Journal of Science, Technology, Engineering and Management - A VTU Publication, 1(2), 27-31. Retrieved from http://ijesm.vtu.ac.in/index.php/IJESM/article/view/232