IoT Data Storage in the Cloud: A Case Study in Human Biometeorology

Brunno Vanelli[1], A.R. Pinto[1], Madalena P. da Silva[1], M.A.R. Dantas[1], M. Fazio[2(✉)], A. Celesti[2], and M. Villari[2,3]

[1] Federal University of Santa Catarina, Florianópolis, Santa Catarina, Brazil
mario.dantas@ufsc.br
[2] University of Messina, Messina, Italy
{mfazio,acelesti,mvillari}@unime.it
[3] IRCCS Centro Neurolesi "Bonino Pulejo", Messina, Italy

Abstract. The IoT (Internet of Things) has emerged to increase the potentiality of pervasive monitoring devices. However, the implementation and integration of IoT devices, data storage and the development of applications are still considered challenging. This paper presents an infrastructure for aggregating and storing data from different sources from IoT devices to the cloud. In order to evaluate the infrastructure regarding the quality in storage, it has been implemented and verified in an AAL (Ambient Assisted Living) case scenario, the main application being Human Biometeorology. The evaluation of metrics related to sending, receiving and storing data demonstrate that the experimental environment is completely reliable and appropriate for the case study in question.

Keywords: AAL · Cloud computing · Human biometeorology · IoT

1 Introduction

The emergence of WSNs (Wireless Sensor Networks) enabled the pervasive monitoring of environments. However, the main weakness of the WSNs is that communications are restricted to the monitoring site (due to short-range radios and energy constraints). The necessary modifications to the WSNs in order to actually be introduced on a large scale in the IT industry (Information Technology) is to connect them to the Internet and extend their limited computation and storage capabilities. Thus, the IoT (Internet of Things) has emerged to fill this gap and provide interconnected devices able to interact with the environment [1].

The implementation and integration of IoT devices, data storage and the development of applications is very challenging. This paper presents an infrastructure for aggregating and storing data collected from different IoT devices into the cloud. It make use of consolidated technologies, such as ZigBee to interconnect monitoring devices, and Azure, to implement a NoSQL Database (DB) into the cloud.

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2018
A. Longo et al. (Eds.): IISSC 2017/CN4IoT 2017, LNICST 189, pp. 253–262, 2018.
https://doi.org/10.1007/978-3-319-67636-4_26

In order to evaluate the proposed infrastructure and the quality of the storage service in, the paper deals with an AAL (Ambient Assisted Living) scenario and, in particular, it addresses an Human Biometeorology application as use case. The evaluation of metrics related to sending, receiving and storing data demonstrate that the experimental environment is completely reliable and appropriate for the case study in question.

This paper is organized as follows: Sect. 2 introduces the reference scenario and the motivations at the base of this work. Section 3 discusses the state of the art on the main topic related to our work. Section 4 presents the architecture we propose and related technologies. We describe our evaluation results in Sect. 5. Finally, Sect. 6 presents our conclusions and future work.

2 Reference Scenario and Motivations

This section presents the motivational scenario, the experimental environment and the results of evaluation of infrastructure proposal. The hardware, software and technologies used in the experiments, are shown in Fig. 1.

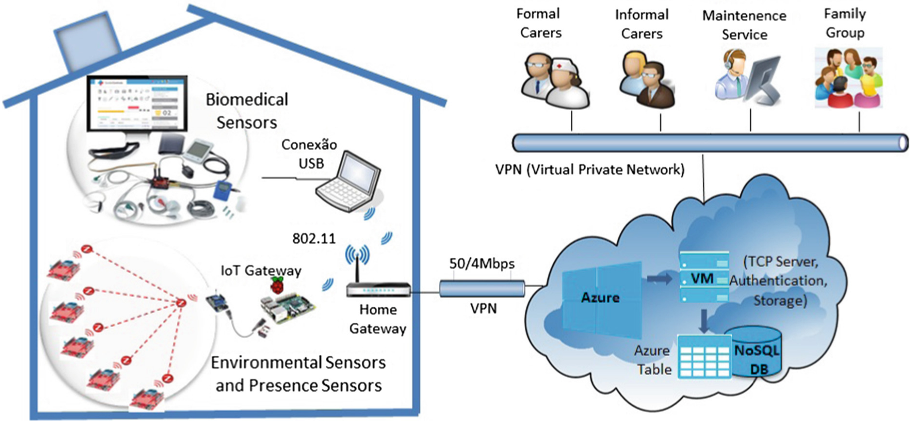


Fig. 1. Experimental infrastructure

The Human Biometeorology is the science that studies the impact of atmospheric influence on health and well-being of humans. There are many proposals in the literature that correlate weather conditions with health, including thermal comfort [2] morbidities [3] mortality [3,4], fetal health [5], among many others. The vast majority of the proposed uses data from the health information providers and climate to correlate and make inferences about the impacts of climate variables with some morbidities of a country, region or group of people [6].

Although the aforementioned research to generate indicators for the management of health, they need to be refined, since there is difference in the accuracy of the values of the meteorological variables read in external versus internal environment [3]. As people live most of the time indoors and many of them are likely to trigger some kind of respiratory disease, it is strongly recommended to monitor the weather conditions in the internal environment. The values of the meteorological variables associated with the user context (i.e. clinical status, patient/family history) become valuable indicators for decision-making by health care providers.

Using technology to support the Human Biometeorology, this article implements the IoT to monitor environmental conditions (dry bulb temperature, dew point temperature, relative humidity, light), the patient's clinical conditions (biomedical signals) and user detection in rooms of an AAL (Ambient Assisted Living) ubiquitous. In the experiments, it was designed a scenario where the patient is remotely assisted by health caregivers and the devices often send data to the cloud relating to biomedical signals, environmental conditions and presence.

Data from the AAL are stored in the cloud and can be consumed by third party applications (formal caregiver/informal, maintenance and family group). The data collected on the case scenario is often crucial for proper monitoring of the patient, hence it should offer reliability and quality in the storage infrastructure proposal.

3 Related Work

The search for related work was conducted on two main topics, that are (1) the adoption of ubiquitous computing to monitor environmental conditions and correlate the meteorological variables with human biometeorology, and (2) data storage solutions for IoT data into the Cloud.

About the first topic, we noticed that many monitoring biomedical signals using body sensor networks and monitoring elderly activities in AAL environments use the ZigBee technology. Indeed, ZigBee presents good performance in monitoring the ambient air quality in order to improve and support the users' health [8,9,11–13]. For these reasons, we adopted ZigBee in our experimentation (as we will discuss later).

About the second topic, we noticed a great interest of the scientific community in designing new solutions for IoT and cloud integration. This paper [14] presents a two-layer architecture based on a hybrid storage system able to support a Platform as a Service (PaaS) federated Cloud scenario. Generalized architectures which use Cloud computing and Big Data for effective storage and analysis of data generated are discussed in [10,15]. This paper proposes [16] a parallel storage algorithm for the classification of data The experiment shows that it classifies the original heterogeneous data flow according to the data type to realize parallel processing, which greatly improves the storage and access efficiency. In this paper [17], the two technologies, cloud computing and IoT, are introduced and analyzed. Then, an intelligent storage management system is designed combining of cloud computing and IoT. The designed system is divided into four layers: perception layer, network layer, service layer, and application layer. And the system's function modules and database design are also described. The system processes stronger applicability and expansion functions, and all of them can be extendedly applied to other intelligent management systems based on cloud calculating and IoT. Our solution is mainly focused on a storage service for IoT data exploiting consolidated technologies, such as ZigBee and Microsoft Azure.

4 Storage Service in the Cloud for IoT Data Management

Figure 2 presents the reference architecture for the IoT data storage service in the cloud. The architecture is composed of three layers, described below.

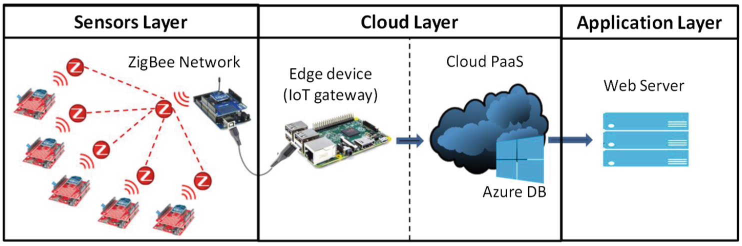


Fig. 2. Reference architecture for IoT data storage service in the cloud

The **Sensors Layer** is responsible for perceiving the environment, collect and transmit the data to the IoT gateway. In our implementation, at this layer we exploit ZigBee networks, because they allow the standalone and scalar configuration, and provide a low-cost and low-power solution. However, any other module or physical medium can be used, as long the devices all agree upon how the data will be handed to the gateway. In particular, each ZigBee network is configured with star topology, where the master node coordinates the process of sending slaves through requests. The slaves collect data from sensors (temperature, humidity, light intensity and presence) according to the master's requests.

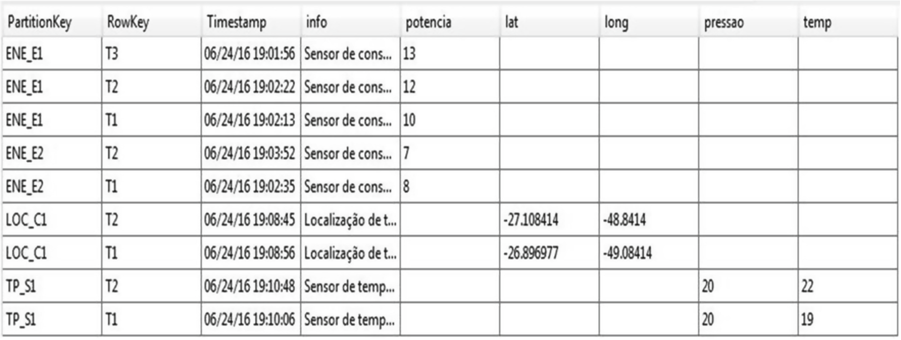
The **Cloud Layer** implements remote resource over the Internet and in cloud datacenters. It includes an *edge device* (e.g., IoT gateway), whose function is to convert data received from the **Sensor Layer** (temperature, humidity, presence and brightness) in data suitable for the **Application Layer**. The Cloud Platform as a Service (PaaS) provides tools for data storage and retrieval over cloud datacenters, in order to benefit of the main advantages of cloud computing in terms of scalability and elasticity. In particular, data are stored in a NoSQL database that is Azure DB.

The **Application Layer** implements the business applications necessary to manage and process AAL data. At this stage of our work, we exploit a web server at the application layer that processes requests for data storage and retrieval.

To store and provide fast retrieval of information generated by the various devices connected to the cloud, it is proposed a storage mechanism based on NoSQL operating architecture, available on Azure called Azure Tables. Unlike storage systems based on relational paradigm, NoSQL databases have as one of its most notable features the non-relational data schema, often allowing more flexible data storage. The Azure tables allow the creation of key-pair tuples (records) with different number of attributes (columns) to be stored in the same entity. Thus, the Azure Tables is able to offer flexible and low-cost storage and efficient searches.

Basic Scheme for Data Storage.
The Azure Table imposes some restrictions on the indexing of stored data. In short, an entity must have two indexed fields: *PartitionKey* and *RowKey*. Additionally, the Azure Table automatically generates a timestamp indicating the last date that the entity was created or changed, for control purposes. The indices of the entities are created using the combination of the PartitionKey with the RowKey, which must be unique across the table. The PartitionKey indicates the partition in the table, and other fields like RowKey are often used to refine the search. Each entity can have up to 255 properties (including the three required) and can store user information such as data strings and integers. According to the manufacturer of the product, the retrieval of information is facilitated when, in the schema definition, the primary information for the search are aimed at these fields because their indexing is already automatically using the combination PartitionKey/RowKey. Searching other fields is possible, but may have some limitations due to the flexibility of the scheme, and is computationally more expensive. On this basis, we defined the basic scheme for storing data of different devices based on a single entity and using the device identifier information, device type and date of the event, to compose the entity's records keys. Thus, each device will be required to provide this information to the composition of the keys and will be free to store in each record the number of attributes you need (e.g., a temperature and humidity sensor will send two pieces of information to each reading, while a location sensor can send, for example, latitude, longitude and altitude). Table 1 shows an example with data on the hypothetical scheme adopted. The PartitionKey field will store the device type (in the example can be: ENE LOC or TP) concatenated with the device identifier (E1, E2, C1, C2, S1). This approach facilitates the retrieval of data from specific sensors. For instance, a query could be made to get all the data related to the ENE sensors, or refine the search to all ENE_E1 sensors. The field RowKey will contain temporal information about the event. The mandatory field Timestamp is an internal information system, and stores the time of the last recording information in the entity. This field can be used for reading but cannot be changed. For this reason, it was decided to store the event occurrence time in the RowKey field, since the time of occurrence of the event and the arrival time to the storage system in the cloud can be distinguished due to delays, in both the network and the queue storage server, or even batch updates. This way, when retrieving data, the query could specify both the sensors and timespan required for the application.

Table 1. Data schema adopted



PartitionKey	RowKey	Timestamp	info	potencia	lat	long	pressao	temp
ENE_E1	T3	06/24/16 19:01:56	Sensor de cons...	13				
ENE_E1	T2	06/24/16 19:02:22	Sensor de cons...	12				
ENE_E1	T1	06/24/16 19:02:13	Sensor de cons...	10				
ENE_E2	T2	06/24/16 19:03:52	Sensor de cons...	7				
ENE_E2	T1	06/24/16 19:02:35	Sensor de cons...	8				
LOC_C1	T2	06/24/16 19:08:45	Localização de t...		-27.108414	-48.8414		
LOC_C1	T1	06/24/16 19:08:56	Localização de t...		-26.896977	-49.08414		
TP_S1	T2	06/24/16 19:10:48	Sensor de temp...				20	22
TP_S1	T1	06/24/16 19:10:06	Sensor de temp...				20	19

5 Experimental Results

In this section, we present the experimentation we performed, providing details on our implementation of the IoT storage service on the cloud, and discussing evaluation results.

5.1 Implementation Details

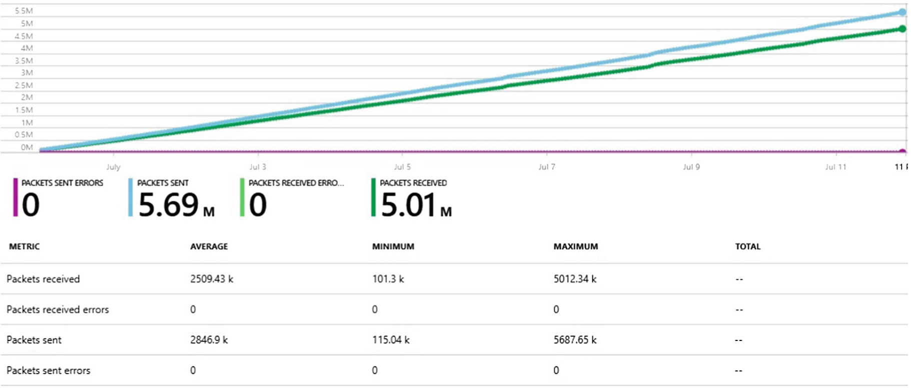
The AAL (Ambient Assisted Living) comprises the ZigBee sensor network and bio-medical sensors. In order to send data to the cloud, it was used a Home Gateway, a TP-Link with Link Internet an Internet link of 50 Mbps/4 Mbps to guarantee access the cloud. TheZigBee network was configured to standard 802.15.4 with star topology consisting of 12 slave nodes and a coordinator node. The slave nodes are composed of 6 DHT11 sensor - humidity and temperature, 3 LDR sensor - light and 3 PIR sensor - presence. The ZigBee network is composed by 12 modules XBee Antenna 1Mw Serie 1. Each module is connected to aXBee shield, which in turn is embedded in the Arduino Uno board.

In the experiment, we used sensors for pulse and oxygen in blood, body temperature, blood pressure,airflow and electrocardiogram sensor (EGC). The collection of data from the sensors was done through the open-source Arduino Software (IDE) - ARDUINO 1.0.6. The data are captured and transmitted via serial communication of the user terminal to the Home Gateway, this in turn sends the data to the web server for storage in the cloud. The frequency of data transmission depends on the type of sensing, that is, pulse and oxygen in blood, body temperature – every 5 min; bloodpressure – every 2 h; airflow and electrocardiogram sensor (EGC) – continuously for periods of set times).

After establishing the connection, the host (i.e., IoT gateway) needs to authenticate to the web server. The authentication process assigns the credentials of the host and grants permission for the storage table(s). To store the data, each host can invoke the storage functions. For each data type sensed, there is a storage function and invocation, and each host must pass the right parameters according to the type of data sensed. After this process, the script automatically sends a data storage request to the respective table in the database. Considering the organization of the data schema defined in Sect. 4, searches by device type, device identifier, or time will be easier and there will be a better use of Azure Table mandatory keys, since these fields carry the most used information such as filter in the consultations to be held. Authentication services and storage were made in the Azure using a virtual machine, standard DS1 v2 (1 core, 3.5 GB memory), with Linux operating system. For this scenario have been implemented some tables, the main ones being the tables for authentication of hosts, storage environmental conditions and storage of biomedical signals.

5.2 Evaluation Results

We evaluated received packets from the IoT Gateway at the web server, and sent packets from the server for the Azure Storage platform. In Fig. 3, it is possible to see that all packets have been received and sent without error. This is justified by the TCP/IP protocol reliability in the exchange of messages.



PACKETS SENT ERRORS
0

PACKETS SENT
5.69 M

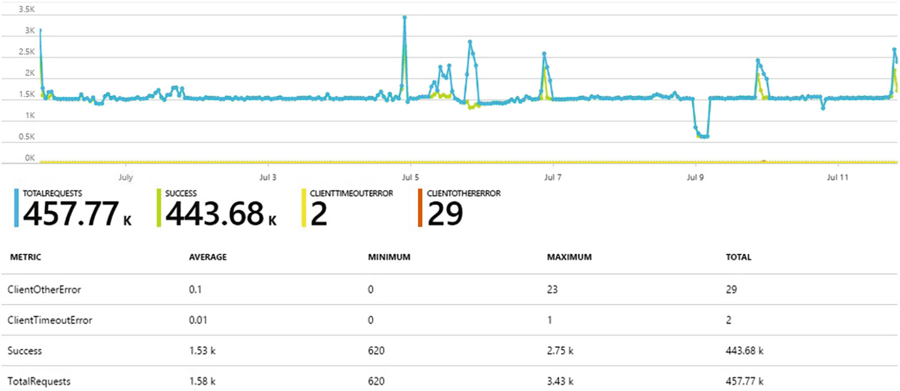
PACKETS RECEIVED ERRO...
0

PACKETS RECEIVED
5.01 M

METRIC	AVERAGE	MINIMUM	MAXIMUM	TOTAL
Packets received	2509.43 k	101.3 k	5012.34 k	--
Packets received errors	0	0	0	--
Packets sent	2846.9 k	115.04 k	5687.65 k	--
Packets sent errors	0	0	0	--

Fig. 3. Monitoring transmission data

Figure 4 illustrates the monitoring Azure Tables metrics: *TotalRequests, Success, ClientOtherError and ClientTimeoutError*. The *TotalRequests* metric summarizes the number of requests made to the storage service. This number includes successful and failed requests and requests that generated errors. The *Success* metric indicates the number of successful requests to the storage service. The *ClientOtherError* metric monitors authenticated requests that failed as expected. This error can represent many of status codes and HTTP 300–400 level conditions as NotFound and ResourceAlreadyExists. The *ClientTimeoutError* metric monitors authenticated requests with time limits that returned an HTTP status code 500. If the client network timeout or request timeout is set to a value lower than expected by the storage service, this will be expected



TOTALREQUESTS 457.77 K

SUCCESS 443.68 K

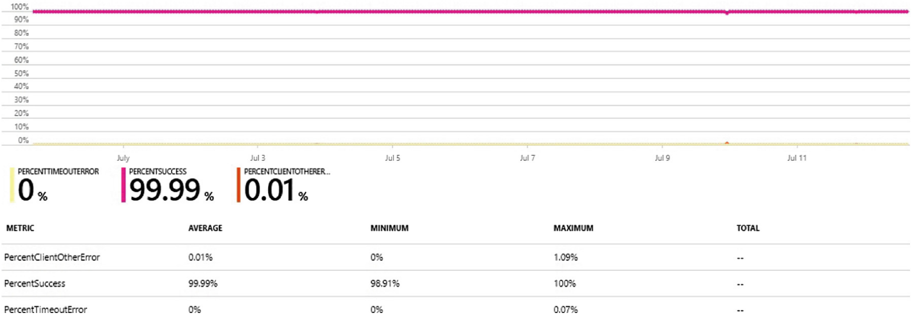
CLIENTTIMEOUTERROR 2

CLIENTOTHERERROR 29

METRIC	AVERAGE	MINIMUM	MAXIMUM	TOTAL
ClientOtherError	0.1	0	23	29
ClientTimeoutError	0.01	0	1	2
Success	1.53 k	620	2.75 k	443.68 k
TotalRequests	1.58 k	620	3.43 k	457.77 k

Fig. 4. Metrics requests for data storage service

time limit. Otherwise, it will be reported as a ServerTimeoutError. Through Fig. 7 can be identified, insignificant, but existing errors by timeout and others. Despite the small difference between total requests (457.77 K) made the Azure tables and the number of successful requests (443.68 K), the average success rate in the tables data storage was 99.99% (Fig. 5).



PERCENTTIMEOUTERROR 0 %

PERCENTSUCCESS 99.99 %

PERCENTCLIENTOTHERER... 0.01 %

METRIC	AVERAGE	MINIMUM	MAXIMUM	TOTAL
PercentClientOtherError	0.01%	0%	1.09%	--
PercentSuccess	99.99%	98.91%	100%	--
PercentTimeoutError	0%	0%	0.07%	--

Fig. 5. Percentages of carried requests the data storage service

Consistent with the metrics monitoring Fig. 4, Fig. 5 shows the metrics with the percentage of success and errors of requests to the storage service. In the monitored period, the minimum percentage of successful requests came in 98.91%, maximum of 100% and average of 99.99%. The percentage of requests that failed with a timeout error, got maximum value of 0.07%. This number includes the client time and server. The percentage of requests that have failed with a ClientOtherError was a maximum of 1.09%.

Figure 6 shows the latency (in milliseconds) of successful requests made to the storage service. This amount includes the processing time required in the Azure storage to read the request, send the response and receive the confirmation response.

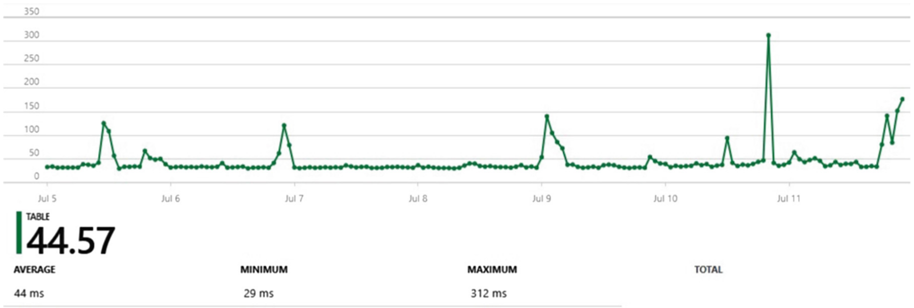


Fig. 6. Time in well successful requests to storage service

6 Conclusions

This paper presented an infrastructure for IoT data gathering and storage service in a NoSQL cloud DB. To evaluate the proposed solution, we implemented the system considering an AAL reference scenario. The setting was applied to human biometeorology where a patient, assisted remotely, frequently sends environmental data, presence and biomedical signals to the cloud. The available data in the cloud can be consumed by third party applications (health caregivers, family members, equipment maintenance operator or the user himself). However, in order to achieve the desired behavior for monitoring applications, it is necessary to verify both the quality in transmission and in data storage. The storage service is based on Azure.

To evaluate the quality, several metrics were selected for the purpose of showing the number and percentage of successful requests to the storage service as well as possible errors and the response time in storage operations carried out successfully. The results show that the experimental environment is reliable and appropriate for the considered case of study.

As future proposals, we intend to scale the AAL equipments, implement new functions in the storage service and work with machine learning, to support human analysis by health caregiver about persisted data.

Acknowledgements. We would like to thanks to Microsoft Research for Azure Award.

References

1. Botta, A., de Donato, W., Persico, V., Pescapé, A.: Integration of cloud computing and internet of things: a survey. Future Gener. Comput. Syst. **56**, 684–700 (2016)
2. Thom, E.C.: The discomfort index. Weatherwise **12**, 57–60 (1959)
3. Quinn, A., Tamerius, J.D., Perzanowski, M., Jacobson, J.S., Goldstein, I., Acosta, L., Shaman, J.: Predicting indoor heat exposure risk during extreme heat events. Sci. Total Environ. **490**, 686–693 (2014)

4. Zhang, K., Li, Y., Schwartz, J.D., O'Neill, M.S.: What weather variables are important in predicting heat-related mortality? A new application of statistical learning methods. Environ. Res. **132**, 350–359 (2014)
5. Ngo, N.S., Horton, R.M.: Climate change and fetal health: the impacts of exposure to extreme temperatures in New York City. Environ. Res. **144**(Pt A), 158–164 (2016)
6. Azevedo, J.V.V., Santos, A.C., Alves, T.L.B., Azevedo, P.V., Olinda, R.A.: Influência do Clima na Incidência de Infecção Respiratória Aguda em Crianças nos Municípios de Campina Grande e Monteiro, Paraíba, Brasil. Revista Brasileira de Meteorologia **30**(4), 467–477 (2015)
7. Yang, C.-T., Liao, C.-J., Liu, J.-C., Den, W., Chou, Y.-C., Tsai, J.-J.: Construction and application of an intelligent air quality monitoring system for healthcare environment. J. Med. Syst. **38**(2), 15 (2014)
8. Sun, F.M., Fang, Z., Zhao, Z., Xu, Z.H., Tan, J., Chen, D.L., Du, L.D., Qian, Y.M., Hui, H.Y., Tian, L.L.: A wireless ZigBee router with P-H-T sensing for health monitoring. In: IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCom 2013, art. no. 682338, pp. 1773–1778 (2013)
9. Nam, J.-W., Kim, H.-T., Min, B.-B., Kim, K.-H., Kim, G.-S., Kim, J.-C.: Ventilation control of subway station using USN environmental sensor monitoring system. In: International Conference on Control, Automation and Systems, art. 6106440, pp. 305–308 (2011)
10. Fazio, M., Celesti, A., Puliafito, A., Villari, M.: Big data storage in the cloud for smart environment monitoring. Procedia Comput. Sci. **52**(2015), 500–506 (2015)
11. Jayakumar, D., Omana, J., Sivakumar, M., Senthil, B.: A safe guard system for mine workers using wireless sensor networks. Int. J. Appl. Eng. Res. **10**(8), 21429–21441 (2015)
12. Sung, W.-T., Chen, J.-H., Wang, H.-C.: Wisdom health care environment systems for monitoring and automated control via RBF function. Appl. Mech. Mater. **157–158**, 315–318 (2012)
13. Li, H., Zhao, L., Ling, P.: Wireless control of residential HVAC systems for energy efficient and comfortable homes. ASHRAE Trans. **116**(PART 2), 355–367 (2010)
14. Fazio, M., Celesti, A., Villari, M., Puliafito, A.: The need of a hybrid storage approach for IoT in PaaS cloud federation. In: 28th International Conference on Advanced Information Networking and Applications Workshops, pp. 779–784 (2014)
15. Behera, R.K., Gupta, S., Gautam, A.: Big-data empowered cloud centric internet of things. In: International Conference on Man and Machine Interfacing, pp. 1–5 (2015)
16. Yan, Z.: The application of mobile cloud in heterogeneous data storage in web of things system. In: 7th International Conference on Intelligent Computation Technology and Automation, pp. 773–776 (2014)
17. Kang, J., Yin, S., Meng, W.: An Intelligent storage management system based on cloud computing and internet of things. In: Patnaik, S., Li, X. (eds.) Proceedings of International Conference on Computer Science and Information Technology. AISC, vol. 255, pp. 499–505. Springer, New Delhi (2014). doi:10.1007/978-81-322-1759-6_57