Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 92 (2016) 583 – 588

Procedia Computer Science

2nd International Conference on Intelligent Computing, Communication & Convergence

(ICCC-2016)

Srikanta Patnaik, Editor in Chief

Conference Organized by Interscience Institute of Management and Technology

Bhubaneswar, Odisha, India

Design of IoT Systems and Analytics in the context of Smart City initiatives in India

Praveen Vijai[a*], Bagavathi Sivakumar P.[b]

[a,b] *Dept of Computer Science and Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, Amrita University, India*

Abstract

The rapid growth of population and industrialization has paved way for the use of technologies like the Internet of Things which gave rise to the concept of smart cities. India as a developing country has a great prospect in developing technologies to make the cities smart. As urbanization occurs the demand for resources and efficient servicing will increase. To achieve this in a smart and efficient way, connected device (IoT) could be used. The possible design of an IoT system based on surveys performed on similar smart solutions implemented has been discussed in this paper. Urbanization and population growth has led to higher demand for resources like water which are of scarce. There is a keen interest from the organizations and government to make proper usage of water. The same can be achieved by proper monitoring and management of water distribution systems. The paper discusses the use of Machine learning techniques to smart city management aspects like smart water management which include water demand forecasting, water quality monitoring and anomaly detection.

Peer-review under responsibility of the Organizing Committee of ICCC 2016

Keywords: Smart City; Internet of Things; Machine Learning; Smart Water Management; Forecasting; Anomaly Detection.

Nomenclature

A praveenvijai90@gmail.com (Praveen Vijai)
B pbsk@cb.amrita.edu (Bagavathi Sivakumar P.)

Peer-review under responsibility of the Organizing Committee of ICCC 2016
doi:10.1016/j.procs.2016.07.386

1. Introduction

India is a rapidly developing country with a high population and growth rate. India accounts for almost 17.51% of the world population. The last decade has seen a dynamic change of rural population to the urban area. It is expected that almost 48% of the country's population would be living in urban area by 2030 which just accounts 33% of the population in 2015 [1]. Due to the growth, there will be a need to address some major challenges related to resources and public services. We are also witnessing immense advancements in networking, software and hardware technology facilitating millions of smart devices and objects to be connected to the Internet. The new technologies will pave way towards generation of smart cities. A Smart City [2] uses the various information and communication technologies to reduce the costs and consumption of resources so as to enhance the well-being of citizens.

Urbanization has its fair share of advantages and disadvantages. Urbanization acts as the driving force for economies of most cities, but there are also concerns of urbanization that impact on the environment, transportation and health issues. Environmental issues include increase in demand of water usage, decline in the air and water quality, decrease in drainage, increase in flood frequency and increase in the cost of city administration [3]. This has led to need for crafty urbanization which inturn has increased the demand for smart cities in India. Key "smart" sectors include health care, transport, energy, water and waste management. Fig. 1 shows some of the possible pillars for the smart city and the applications related to it. The future internet technologies and the presence of Internet of Thing (IoT) [4] is of great use in the smart city situations. Initially the terms IoT was used to indicate uniquely identifiable connected objects with radio-frequency identification (RFID) technology. Later on, the RFID tags, sensors/actuators and communication technologies were brought as the building blocks for the IoT. A variety of physical and mechanical objects around us can be connected to internet and allowed to communicate and cooperate with each other to attain some goals.

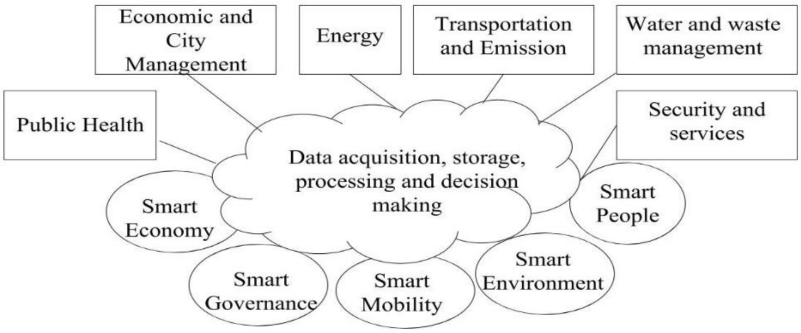

Fig. 1. Various Smart City applications and pillars [5]

As smart cities have paved way for talk on better utilization of resources, and also the exponential growth of population in the cities has led to a high demand on resources like water and proper usage of the same is in the spotlight. Effective water usage is very context specific, since it depends on many factors like the climate, resources, users etc. This means that the water saving solutions derived for one organization may not work for another. The growth of IoT has led to use of sensors to monitor different aspects on the usage of water and make analysis on these data.

The remainder of the paper is organized as follows. Section 2 of the paper mentions about the Indian government plan on setting up smart cities. Section 3 discuss on Internet of Things and various constrains of implementing an IoT system in India. Section 4 provides a brief discussions on the machine learning methods used in the water management systems. Finally in Section 5 the conclusion is made.

2. Smart city initiatives in India

The Indian government has realized the need for planning for smart cities and has come up with various schemes for the same. As a beginning step the government has declared probable 100 cities in India that could be developed to Smart Cities mostly based on the population of cities and has also allocated Rs 7,060 crores in 2015 Budget for the same [6]. The distribution of 100 cities is like nine satellite city with a population more than 4 million. Almost 19 cities of the list are the capital of state / Union Territory and 10 cities with religious and tourist importance are also named in the list. Also 37 cities and 25 cities which has a population range of 4-1 million and 1-0.2 million respectively are included in the potential shortlist. The Indian government has also announced $15 billion for the solo development of IoT in the Indian market. Thus there is wide possibility for developing solutions that can help in the process of implementing smart cities in India.

3. Internet of Things (IoT) Systems

As an emerging technology IoT is getting widely accepted all around the world. An IoT [7] can be defined as a network of connected embedded objects or devices, with identifiers, in which a communication using standard communication protocol can happen without human intervention. Reports suggest that the number of internet connected devices has surpassed the human population of the world and this IoT devices act as the building block for Smart cities. As per Gartner research the growth in the smart technology is being depicted in fig.2. As shown, there would a high number of smart homes and commercial building in which the smart power and water management are essential criteria to be met.

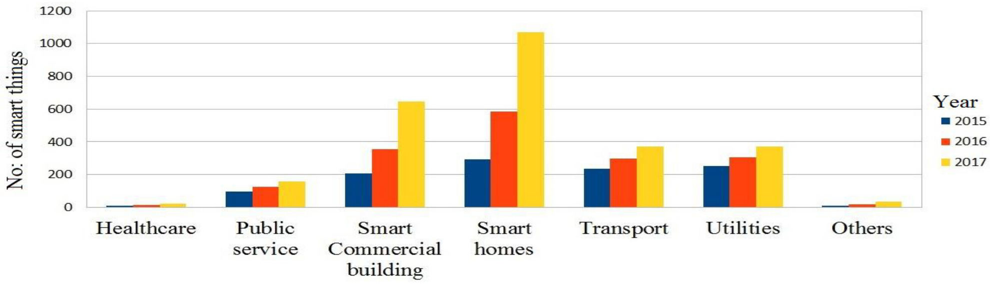

Fig. 2. Gartner forecast on growth of pillars of smart city [8]

The following are recognized as the standard stages in setting up an IoT system [7,9]:

- The sensor that collects the data from the environment (including the identification address of the sensor)
- An application that is used to collect and analyze this data to infer knowledge from it.
- Decision making and transmitting of information to the necessary hubs. Actuators and big data analytics is used for the same.

Based on the study performed on the various IoT systems the typical structure of an IoT solution is realized. Fig.3. shows a framework/architecture for a typical IoT system. An IoT system consists of a set of sensors that measure any change in the environment. This data measured are to be transmitted to a central/decentralized system via. a connectivity like 3G, Bluetooth, Zigbee etc depending on the distance and data speed required. There is also a need for providing steady power supply to the sensor system, the power capacity mainly depends on the connectivity as connections like 3G drain more power than a Bluetooth system, which is to be looked into while choosing a connectivity. Also the security aspect of the hardware and connectivity is to be taken care of. The main requisite for an IoT system is the solution should be usable for everyone and not just an expert. The data received in the cloud system are stored or processed for the discovering patterns and to infer knowledge. The visualization of the data can be performed for easy understanding to the end user and alert systems can be made use to provide appropriate warning.

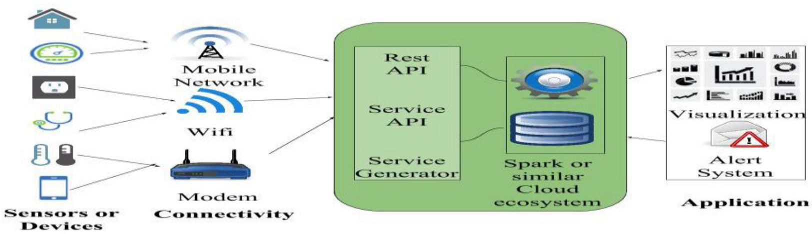

Fig. 3. Framework for an IoT System

4. Machine Learning for Knowledge discovery

As mentioned it is important for the system to make use of the data generated by the IoT inorder to infer knowledge from them that could benefit the consumers and the government. This is where the various machine learning techniques come into play as these methods can be used to extract hidden information from data of IoT. This knowledge can used to improve the performance of the system and the quality of service. The smart water management can be divided mainly into three parts (1) forecasting demand, (2) water quality monitoring and (3) anomaly detection. This section deals with the machine learning technologies used for the same.

4.1. Water demand forecasting

The optimization of water management system can be done cost effectively by using prediction. For a better planning and also to meet up with the demand for a resource it's always essential to plan for the future, by forecasting we can achieve it. The Time series analysis is the most widely used forecasting method, there are many papers that use the techniques. The factors like seasonality, level shift factor and trend are key elements while formulating a time series model. Different models have been developed for making analysis in time series data. The Autoregressive Integrated Moving Average (ARIMA) model comprises of differentiation and integration process. When the seasonality property of time series is reflected then the model is extended to SARIMA. The time series data with auto correlation makes use of the Autoregressive Fractionally Integrated Moving Average (ARFIMA) [14]. The Holt-Winters (HW) method account both seasonality and trend and is an extension of exponential smoothing. The ARIMA is widely used in linear prediction as it is simple and flexible but the model has poor performance for nonlinear problems. However the Radial Basis Function (RBF) could overcome this problem as it handles nonlinear relations. Li Jie et al.[19] has used a model which is a combination of both methods to forecast the water demand.

Seok et.al.[11] discusses on Error percentage Correction(EPC) algorithm which uses weighted hourly usage and error feedback to forecast the demand hourly. Genetic Programming (GP) with Extended Kalman filter (EKM) is used by Fagianio et al.[12] where GT selects the best model and EKF predict states. Magiera [15] uses Bayesian Network for forecasting and also proposed a new accuracy measurement called Discrete Mean Absolute Forecasting Error (DMAE). The Dynamic Gaussian Bayes Network (DGBN) that select random variable at discrete time and uses linear conditional Gaussian distribution is implemented by Wojciech et al.[13]. Ji et al.[16] tuned the Least Square Support Vector Machine (LSSVM) with enhanced teaching learning based optimization (TLBO) to obtain a better regression accuracy. Bakker et al. implemented a model which works in three phase, first the average water usage for 48hrs is forecasted; then the normal demand for 15 min time steps and finally the abnormal demand for 15min time steps. The various work done in this regard are summarized and is shown in table. I.

From the surveys performed in various forecasting models of water demand we have made the following inference.

- Input parameters: There are fair number of papers in which the model just makes use of the water demand as a single input and still able to make a fairly accurate forecast. Other models make use of the measured weather information like the temperature, humidity, rainfall etc as the input along with water demand.
- Time step: The time step denotes the time difference in which the consumption on water have been recorded. Time step for a model is decided by the purpose for which the model is being used. The model can be for hourly basis, daily or monthly. The smallest time step found in related papers is one hour, which is a large gap as it can't describe all the variations in the demand.
- Demand patterns: Most of the models use different patterns for different days of week: one for weekdays, one for Saturday's and another for Sunday's. Also there are papers in which national holidays are taken into consideration.
- Model recalibration: Almost all the models are formulated on some historical data and so the model needs to be recalibrated with new datasets when implementing in real life situations.

Table I. Summary of papers on forecasting water demand

Paper	Dataset	Samples	Period	Techniques	Evaluation criteria	Performance
[11]	Sugi Water Treatment plant	8790	Hourly	EPC	MAPE	9.38
[12]	AMPds	10,51,200	1hr, 6hr, 12hr and 24hr	GP and EKF	NMSE	0.083
					R^2	0.917
[13]	Sosnowiec Water Distribution Data	2386	Daily	DGBN	MAPE	7.6
[14]	Sosnowiec Water Distribution Data	-	-	SARIMA	MAPE	37.29
				ARFIMA		36.47
				HW		48.10
[15]	Sosnowiec Water Distribution Data	-	-	BN	DMAE	212.28
[16]	Shanghai Water Supply System	8736	Hourly	TLBO and LSSVM	MRE	0.41
[17]	Syarikat Air Johor(SAJ)	204	Monthly	EMD and LSSVM	MAE	1.6485
					RMSE	2.0021
					R	0.6819
[18]	Netherlands water usage	2,10,336	24hr and 48hr	Adaptive	MAPE	2.91
					RRMSE	4.35
					R^2	0.75

4.2. Water quality and Anomaly Detection

It is essential to monitor the quality of a life sustaining resource like water. The machine learning techniques like Artificial Neural Network (ANN) [20] and fuzzy systems [21] are widely used for this purpose. The main sensing parameters that are measured to analyse the water quality are level of turbidity, chlorine, ORP, Nitrates, pH, conductivity and temperature. The techniques like sliding window are used by Oliker et al. [22] to find the similarity with other windows so as to detect the event causing it. Also hybrid model are used for prediction of the water quality like ANN and markov chain [23] is used to predict the oxygen content in water also models like SVM [24] are used to provide early warring.

The anomalies are those data which don't follow the usual pattern of the data, such data are also called outlier. By statistical techniques the outlier data are those which are far from the mean which is discussed in [25, 26]. But the problem with the method it assumes the data follows a distribution which is not the case for high dimension data. Another method by which anomaly can be detected by nearest neighbor where the distance to every data point is measured and those above a threshold are labelled outlier [27]. But the problem of the method is it computationally complex. Also clustering based approaches can be made use for anomaly detection, in which normal data will be large clusters whereas anomaly will be small clusters.

5. Conclusion

In this paper we have proposed a general framework for IoT systems based on the analysis done on various Internet of Things solutions in the context of smart city. Further the paper also discusses on the Machine Learning techniques used in smart water management. The smart water management mainly deals with forecasting of the water demand so as the authorities can plan for future requirements. It also deals with quality monitoring of water and anomaly detection. A very brief discussion of the same is done in the paper. As summary presented shows there are numerous models used for forecasting of water demand. However despite certain models performing well, the systems surveyed in this paper still requires further improvements. Thus there is a huge potential of using IoT systems and analytics in the design of smart cities in India. The design of such a system and its applications have been discussed in brief.

References

1. Worldometer Population: http://www.worldometers.info/world-population/india-population/
2. A. Domingo, B. Bellalta, M. Palacin, M. Oliver, and E. Almirall, “Public open sensor data: Revolutionizing smart cities,” IEEE Technology and Society Magazine, vol. 32, no. 4, pp. 50–56, 2013.
3. Basawaraju, R., et al. (2011). “Analysis of the impact of urban sprawl in altering the land-use, land-cover pattern of Raichur City, India, using geospatial technologies.” J. Geogr. Region. Plann., 4(8), 455–462.
4. L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.
5. Z. Khan, A. Anjum, K. Soomro, and M. A. Tahir, “Towards cloud based big data analytics for smart future cities,” Journal of Cloud Computing, vol. 4, no. 1, 2015.
6. Mathews N., India’s ruling party moves forward on ’100 smart cities’ plan, ENR (Engineering News-Record) 274 (14).
7. J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of things (iot): A vision, architectural elements, and future directions,” Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, 2013.
8. Gartner Trend Prediction: http://www.gartner.com/technology/topics/trends.jsp
9. D. Bandyopadhyay and J. Sen, “Internet of things: Applications and challenges in technology and standardization,” Wireless personal Communications, vol. 58, no. 1, pp. 49–69, 2011.
10. Xiang, P. (2012). Design of smart home system based on the technology of internet of things. *Research Journal of Applied Sciences, Engineering and Technology, 4*(14), 2236-2240.
11. Seok, J. , Kim, J. , Lee, J., Song, Y. -., & Shin, G. -. (2014). Abnormal data refinement and error percentage correction methods for effective short-term hourly water demand forecasting. *International Journal of Control, Automation and Systems, 12*(6), 1245-1256.
12. Fagiani, M., Squartini, S., Gabrielli, L., Spinsante, S., & Piazza, F. (2015). *Domestic water and natural gas demand forecasting by using heterogeneous data: A preliminary study*
13. Froelich, W. (2015). *Forecasting daily urban water demand using dynamic gaussian bayesian network* doi:10.1007/978-3-319-18422-7_30
14. Froelich, W. (2015). *Dealing with seasonality while forecasting urban water demand* doi:10.1007/978-3-319-19857-6_16
15. Magiera, E., & Froelich, W. (2015). *Application of bayesian networks to the forecasting of daily water demand*
16. Ji, G., Wang, J., Ge, Y., & Liu, H. (2014). Urban water demand forecasting by LS-SVM with tuning based on elitist teaching-learning-based optimization. Paper presented at the *26th Chinese Control and Decision Conference, CCDC 2014,* 3997-4002.
17. Shabri, A., & Samsudin, R. (2015). Empirical mode decomposition-least squares support vector machine based for water demand forecasting. *International Journal of Advances in Soft Computing and its Applications, 7*(2), 38-53. Retrieved from www.scopus.com
18. Bakker, M., Vreeburg, J. H. G., van Schagen, K. M., & Rietveld, L. C. (2013). A fully adaptive forecasting model for short-term drinking water demand. *Environmental Modelling and Software, 48*, 141-151. doi:10.1016/j.envsoft.2013.06.012
19. Jie, L., Qiang, L., Yi, H., Liang, L., Cheng, F., & Zhenzhen, J. (2014). *Application of the fully data-driven combination model for water demand forecasting in the mountainous tourist area* doi:10.1007/978-3-642-54927-4_56
20. Wu, W., Dandy, G. C., & Maier, H. R. (2014). Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. *Environmental Modelling and Software, 54*, 108-127.
21. Kalavrouziotis, I. K., Pedrero, F., & Skarlatos, D. (2015). Water and wastewater quality assessment based on fuzzy modeling for the irrigation of mandarin. *Desalination and Water Treatment,*doi:10.1080/19443994.2015.1110050
22. Li, X., & Song, J. (2015). A new ANN-markov chain methodology for water quality prediction. Paper presented at the *Proceedings of the International Joint Conference on Neural Networks, , 2015-September*doi:10.1109/IJCNN.2015.7280320
23. Oliker, N., Ohar, Z., & Ostfeld, A. (2016). Spatial event classification using simulated water quality data. *Environmental Modelling and Software, 77*, 71-80. doi:10.1016/j.envsoft.2015.11.013
24. Liu, S. -., Xu, L. -., & Li, D. -. (2015). Water quality early-warning model based on support vector machine optimized by rough set algorithm. *Xitong Gongcheng Lilun Yu Shijian/System Engineering Theory and Practice, 35*(6), 1617-1624.
25. Hutton, C., & Kapelan, Z. (2015). Real-time burst detection in water distribution systems using a bayesian demand forecasting methodology. Paper presented at the *Procedia Engineering, , 119*(1) 13-18. doi:10.1016/j.proeng.2015.08.847
26. Gu, S. -., Zhang, H., Xie, X. -., & Liu, Y. (2010). Study on the medium and medium-short term characteristics of anomalies of underground water level in hainan province. *Dizhen Dizhi, 32*(4), 638-646. doi:10.3969/j.issn.0253-4967.2010.04.011
27. Almalawi, A., Yu, X., Tari, Z., Fahad, A., & Khalil, I. (2014). An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems. *Computers and Security, 46*, 94-110. doi:10.1016/j.cose.2014.07.005