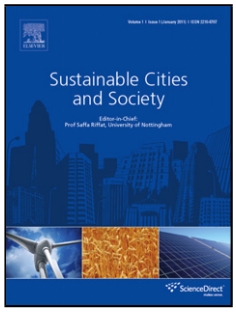
Accepted Manuscript



Title: Exploiting IoT and Big Data Analytics: Defining Smart Digital City using Real-Time Urban Data

Authors: M. Mazhar Rathore, Anand Paul, Won-Hwa Hong, HyunCheol Seo, Imtiaz Awan, Sharjil Saeed

PII: S2210-6707(17)30978-2
DOI: https://doi.org/10.1016/j.scs.2017.12.022
Reference: SCS 894

To appear in:

Received date: 1-8-2017
Revised date: 7-11-2017
Accepted date: 19-12-2017

Please cite this article as: Rathore, M Mazhar., Paul, Anand., Hong, Won-Hwa., Seo, HyunCheol., Awan, Imtiaz., & Saeed, Sharjil., Exploiting IoT and Big Data Analytics: Defining Smart Digital City using Real-Time Urban Data.*Sustainable Cities and Society* https://doi.org/10.1016/j.scs.2017.12.022

Exploiting IoT and Big Data Analytics: Defining Smart Digital City using Real-Time Urban Data

M. Mazhar Rathore[1], Anand Paul[1,*], Won-Hwa Hong[2], HyunCheol Seo[2], Imtiaz Awan[3], Sharjil Saeed[3]
[1]School of Computer Science and Engineering, Kyungpook National University, Daegu, Korea[2]School of Architecture, Civil, Environmental and Energy Engineering, Kyungpook National University, Daegu, Republic of Korea
[3]Department of Computer Science and Information Technology, University of AJK, Pakistan
rathoremazhar@gmail.com, paul.editor@gmail.com*, hongwh@knu.ac.kr, notsools@gmail.com, thisisawan@gmail.com, rajasharjeel@gmail.com

Highlights

Data Generation, Collection, Aggregation, Filtration, Classification, Preprocessing Computing and Decision Making

Impementation of Smart IoT based Digital City using Real-Time Urban Data

Big Data Analytics for City Planning using Hadoop Ecosystem

Big Graph Processing for Traffic Information and Alert

System is Evaluated for its Scalability and Real-time Data Processing

***Abstract*— Integration of all smart systems (such as smart home, smart parking, etc.) and the IoT devices (such as sensors, actuators, and smartphones) in the city can play a vital role to develop the urban services by building their city digital and smarter. However, interconnection of lots of IoT objects to collect urban data over the Internet to launch a smart digital city, effects vast volume of data generation, termed as Big Data. Thus, it is a challenging task to integrate IoT devices and smart systems in order to harvest and process such big amount of real-time city data in an effective manner aimed at creating a Smart Digital City. Therefore, in this paper, we have established an IoT-based Smart City by using Big Data analytics while harvesting real-time data from the city. We used sensors' deployment including sensors at smart home, smart parking, vehicular networking, surveillance, weather and water monitoring system, etc., for real time data collection. The complete system is described by its proposed architecture and implementation prototype using Hadoop ecosystem in a real environment. In addition, the Smart Digital City services are extended by developing the intelligent Smart Transportation System by means of big graph processing to facilitate citizens while providing real-time traffic information and alerts. The proposed system consists of number of stages including data generation and collection, aggregation, filtration, classification, preprocessing, computing, and decision making. The efficiency of the system is extended by applying Big Data processing using Apache Spark over Hadoop. Whereas, the big city graph processing is achieved by using Giraph over Hadoop. The system is practically implemented by taken existing smart systems and IoT devices as city data sources to develop the Smart Digital City. The proposed system is evaluated with respect to efficiency in terms of scalability and real-time data processing.**

***Index Terms*—Smart Digital City, Smart Transportation System, Big Data, Hadoop.**

1. INTRODUCTION

The integration of pervasive and ubiquitous computing changes the style of living. One of the reports notified that in 2050, seventy percent of the world population will live in cities (Jiong et al., 2014). Hence, a rapid growth has been seen in the movement of the population towards cities. Therefore, for urbanization, it is an extreme important aspect to realize the demand for service profiling to improve the competence. It may lead to the modern advancement in the management and control of the city. Presently, few companies developed their IoT platforms for monitoring, aggregating urban data, processing, and planning for urban parameters. For Example, Capillary devices are used to establish a smart cyber society (Ahmad et al, 2016a). In Japan the broadband access offers the communication services from people-to-people, people-to-things, and things-to-things (Srivastava, 2004). Likewise, smart home produced in South Korea allows their residents to control objects remotely (Giroux and Pigo, 2005). The next generation I-Hub of Singapore (Han, 2005) planned to realise the future generation "U" network through a ubiquitous secure network (O'droma and Ganchev, 2010). Hence, it results in the rise of a number of things connected over the internet. Thus, it generates an incredible amount of the diverse and high-speed data, referred as big data.

Big data analytics are very vital for modern science and commercial organisations. This huge amount of data can be generated from various sources, such as sensors, smart systems, online emails, transactions, search queries, audios, videos, social networks, media files, etc (Rathore et al., 2015, Xia et al., 2012). The current studies for smart city development (Kyriazis et al., 2013) and statistical big Data parameters are defined in a limited way (Jianguo and Dingding, 2013; Kleiner et al., 2012; Runze, Dennis, and Bing, 2013). Thus, we cannot fully trust on their measurements. To evident these statistical quantities of the big data sets, someone has to represent them in other suitable forms. Also, it is challenging to aggregate the city data, through the use of cyber physical system (CPS) or IoT devices for users and authority's decisions making to fulfil the demands of inhabitant in order to establish digital smart city. Yet, the systems, mentioned above, work have limitations, as they did not consider the aspect of Big Data generation and its processing. Generally, the data harvesting, aggregation, and analysis are hard to accomplish in such high-speed