IEEE ComSoc
IEEE Communications Society
e-Health Technical Committee
IEEE

e-Health Technical Committee NewsLetter

January-February, 2017

On behalf of the e-Health Technical Committee (TC) of the IEEE Communications Society (ComSoc), we wish all our members a very instructive reading of this letter.
The contribution for this edition is coming from France. The author is sharing his ideas on the provisioning of advanced e-Health services within the field of Cloud Computing. This work is a joint collaboration between researchers in France and Brazil in the area of e-Health.
Members of the e-Health community are invited to contact the author for further information or collaborations.

We also welcome all our members to share their research activities and field experiences through this open newsletter and to open up new opportunities for discussions and collaborations.

Dr. Nada Philip (Kingston University London, UK) &
Dr. Nazim Agoulmine (University of Evry Val d'Essonne, France)

Table of Contents

TOWARDS AIRCARE: A HIGHLY SCALABLE DATA MANAGEMENT SYSTEM FOR MOBILE SENSOR DRIVEN EXPOSURE ANALYSIS TO AIR POLLUTION

Rafiqul Haque, Yehia Taher, and Karine Zeituni
Karine.Zeitouni@uvsq.fr
(Données et Algorithmes pour une Ville Intelligente et Durable (DAVID). Université de Versailles Saint-Quentin-en-Yvelines, France)

ABSTRACT - Air pollution is a long-standing issue, which is at the root of diverse risks to human health. It is responsible for the increase of mortality and morbidity amongst the exposed population. Therefore, it has turned into a global concern and drawn an enormous attention to both profit and nonprofit organizations. Nowadays, most developed countries become concerned by this phenomeno and have established an observation network in order to monitor air quality, mainly in dense urban areas. The advent of Internet of Things (IoT), e.g., sensors, enables to build a community-based observation network, allowing the participant to measure their personal exposure, and to share the sensed air quality measures broadly. However, existing solutions have limitations such as scalability. In this paper, we propose a scalable and reliable system called *AirCare* for air quality and health-effect monitoring. We provide a detailed description of the different components of *AirCare*.

I. INTRODUCTION

Environment pollution has existed for centuries [9]. However, it started to be a significant issue since the industrial revolution in the 19th century. Today, it is one of the greatest problems, which causes grave and irreparable damages to the ecological system [4]. In addition, it poses a wide variety of risks to human health including life-threatening cancerous disease. Therefore, environment pollution has turned into a global concern;

specifically, the experts such as environmentalists and organizations including United Nation (UN) and World Health Organization (WHO) who are heavily concerned about the severe effects of different forms of pollutions.

Environment pollution consists of four different forms: air, water, soil, and sound pollution. Of all, air pollution is the most harmful by far. It is caused by various hazardous gas and particle substances such as carbon monoxide and black carbon, which are found in smoke. Two different types of air pollution are reported in literature [2] [8]: outdoor pollution which primarily results from combustion of fossil fuel by industrial plants or heating facilities, and vehicles; the indoor air pollution, as opposed to the outdoor, which is caused within and around building and structure by volatile organic compounds (VOC). According to WHO, air pollution is carcinogenic, which promotes life threatening health risks. More than 6.5 million people die every year due to air pollution [5]. The International Energy Agency reported that the premature death due to outdoor pollution would rise to 4.5 million by 2040 [4].

Several initiatives have been taken to protect human health from the detrimental effects of air pollution. Information system is used as an enabler to monitor air quality in fixed locations or to estimate pollution level at a regional scale. Indeed, in addition to in-situ field measurements, monitoring ambient air is based on numerical tools to model the spatial distribution of pollution and predict air quality for the coming days. However, this may not be sufficient when one deals with health impacts and the relationships between health impacts and pollution episodes or chronic levels of pollution, because individuals spend most of their time indoors (at home, at work, etc.) or in specific micro-environments along their daily activities and mobility, which are not captured by the approaches of air quality monitoring.

However, with the advent of the Internet of Things (IoT), in particular wearable sensors and mobile devices have led to an enormous change in the environment monitoring system landscape [14]. Over the last few years, several IoT-driven air pollution monitoring and measuring solutions have been developed within several projects (e.g., OpenIoT [6], OpenSense [1], OpenSense II [11], and HackAIR [10]). These solutions perform various tasks including collecting and processing data, performing analysis, and visualization. We investigated the functional and architectural aspects of existing IoT-driven air pollution monitoring systems. Our study reveals that these solutions lack three features: large volume of data handling, physical scalability of the architecture, and usability, which are critical for large-scale deployment of air quality and personal exposure monitoring. Considering these limitations, we propose a highly scalable system called AirCare for managing, analyzing, and visualizing massive-scale air pollution and individual exposure data. This system is under development in the framework of the French multidisciplinary project Polluscope [12]. The objective of Polluscope is to develop a participatory observatory of the surveillance of exposure to air pollution [7]. The detail of this system is discussed in the rest of this paper.

II. THE ARCHITECTURE OF AIRCARE

AirCare is a micro-service oriented distributed system, which comprises a set of diverse decentralized service components. The micro-service paradigm enables building complex solution that is composed of self-contained and autonomous components (applications), which can communicate with each other using language-agnostic APIs (Application Programming Interfaces). It guarantees elasticity of the AirCare system by enabling users to deploy service components with minimal effort and without any complexity. Additionally, services can be turned off without affecting other service components.

The AirCare adopts shared-nothing architectural style. The reasons are two-fold: avoid a single point of contention and increase scalability. Figure 1 shows the end-to-end architecture of AirCare. AirCare relies on hybrid computational model. It supports real-time and batch operations. In addition , the operations are scalable which enables to span operations from a single node to multiple nodes.

A. PHYSICAL COMPONENTS

AirCare system involves three different types of physical components: sensors, mobile devices, and nodes.

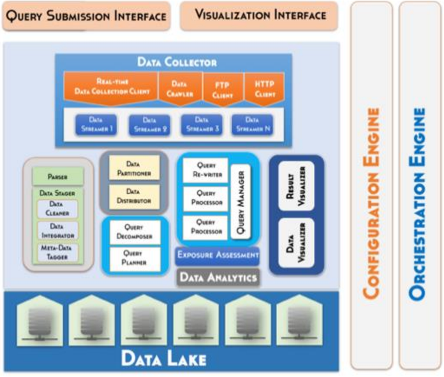

Figure 1: The Architecture of AirCare system

- Sensors are components that detect and monitor the quality of pollutants in air. Within AirCare there are indoor and outdoor air pollution sensors. The indoor sensors are deployed within buildings such as houses to detect VOCs whereas the outdoor sensors are deployed in different locations to measure the value of various pollutant. Furthermore, AirCare system involves health sensors in order to monitor the holder's vital signs.
- Mobile devices (e.g., smartphones) may include various sensors such as accelerometer, global positioning system (GPS), gyroscope, etc. Mobile devices can thereby detect or capture the user's activity. Additionally, they are used as an interface for visualizing the user's exposure level and can be used for communication purposes in case the sensors are not directly connected to the server.
- Nodes are physical or virtual servers, which are used for performing tasks (e.g., enrichment or query processing) and act as a repository for storing data and metadata.

B. CORE FUNCTIONAL COMPONENTS

The AirCare system covers a wide range of functions related to managing, processing, analyzing, and visualizing pollution data.

- Data collector collects data from a list of heterogeneous sources including sensors, legacy databases, and open data. AirCare collects data using two different styles: batch and real-time. The batch style is used since the solution employs opportunistic sensing where sensor devices (or the paired smartphone) store data natively for a certain period and transmit them after some time interval to AirCare's solution framework. Data are collected in real-time as well. AirCare system collects external data specifically open data such as demographic data or weather data from the Web.
- Parser: AirCare's parser can parse different types of data including time series related to environmental, spatial and user's activity and health. It is able to handle both structured and unstructured data models.
- Data Staging component perform several functions. It consolidates (integrates) data coming from a wide variety of sources. Data cleansing refers to removing invalid data and handling missing and inconsistent data. Additionally, the noisy data are dealt with in this phase. Furthermore, it enables tagging meta-information and helps in enriching data by adding external data with data stemming from mainstream (primary) sources.
- Data Transformation component performs the mapping from the received format to the target representation model. This model encompasses the different complementary aspects of the data, including Time series, spatial data and metadata to form a completely novel holistic vision.
- Data Partitioner splits dataset into fixed size data chunks. Partitioning can be performed randomly. However, AirCare supports rule-based data partitioning as well. In this type, datasets are split based on a set of predefined rules.
- The Data Distributor distributes data in the cluster of nodes. Data distributor of the system can perform random distribution of data and policy based distribution. The random distribution relies on the replication policy of the Hadoop Distributed File System (HDFS) [13].
- Query Decomposer splits a complex query with two or more joins into atomic sub-queries. The key

purpose of decomposing queries is to increase parallelization in query processing.

- Query Planner plans an execution of queries. Since data are distributed within and across the cluster of nodes, the planner plays a critical role to optimize queries by decomposing and distributing the queries efficiently within the cluster.
- Query rewriter rewrites queries based on the query plan. The key purpose of re-writing is to send sub-queries to the nodes that store the data; hence, the processing time is reduced.
- Query Processor performs the distributed and parallel execution of sub-queries. It also performs post-processing tasks such as sorting and merging the results.
- Exposure assessment component relies on aggregation functions that compute exposure indicators (cumulative sum, activity intensity, inhalation, etc.) of an individual exposed to indoor and outdoor pollutants.
- Data Analytics will carry out analysis of health risk. It will perform analysis of short-term effects of exposures on health. The analytical engine relies on machine learning and data mining techniques. Additionally, the engine will analyze the effects of pollution to specific types of individuals in particular, the patients who are hospitalized, and will correlate the observations with previous ecological studies.
- Visualizer presents data or outcomes. The visualizer presents raw data stemming from sensor and the results of analytical queries.

C. MANAGEMENT COMPONENTS

The AirCare system integrates a wide range of management components.

- Configuration manager deals with the configuration of components. The binding of service components to build the seamless AirCare system is managed by the configuration manager.
- Orchestrator coordinates or manages the interactions and operations amongst the component.
- Query Manager coordinates the query processing tasks.

AirCare system query submission interface and visualization interface. The former enables submitting queries and the latter is used to visualize data or outcome. The latter provides an access to the system through different channels including cell phones.

II. HOW DOES IT WORK?

The data collection services of AirCare are embedded with the sensors, which push data to the AirCare system. The data collectors fetch data streams and ingest them to data processing components. In some cases, sensor data are collected via FTP and HTTP clients, which assist to fetch data from the server of the service provider and cache them in the disk. We use disk-caching technology for storing data temporarily. In addition, data are collected using Web Crawler (See Figure 1). The data streamer streams data to the processing components. Then the parser parses the data and stores them in the data stage server where they are cleaned and consolidated. For cleaning data efficiently and to ensure the quality of data, users can visualize the raw data and find missing elements easily. The external data are integrated to enrich the datasets n this phase. In addition, the user can tag them with meta-data using a tagging application provided by AirCare.

The partitioner splits datasets into smaller chunks, which are distributed to a data lake, which is a cluster of nodes. Users can query data through the query interface provided by the AirCare system. The visualizer allows presenting the results of all sorts. The exposure assessment of an individual is carried out periodically and a comprehensive report, including a dashboard and a map, is presented through a mobile interface. The results are also accessible via the web interface. The data analytics performs analyses to compute the health risk of individuals. The outcome of the analysis is reported through a visualization interface.

IV. CONCLUSION

Air pollution has turned into a global concern after industrialization was revolutionized. Millions of people die every year due to air pollution. To control air pollution and to save human life, several projects have been initiated. Information technology and advanced technologies such as the Internet of Things (IoT) are used enormously to drive air pollution monitoring systems. In this article, we presented the AirCare system which we envisaged for managing and analyzing massive air pollution data to measure exposure level of healthy individuals and the ones with health issues, and visualize these data. The proposed system adopts hybrid computation model to process data both in real-time and batch operation styles. They can be used based on the requirement. We explained a range of components included in the system, which we plan to develop in the near future. We also explained how the system will work in the real-world. The imminent future task we plan is developing the components of the AirCare system in order to make it fully functional for the users.

ACKNOWLEDGMENT

This work is supported by the grant ANR-15-CE22-0018 POLLUSCOPE of the French National Research Agency (ANR).

REFERENCES

[1] K. Aberer, S. S. Chakraborty, D., Martinoli, A. G. Barrenetxea, B. Falting, and L. Thiele, "OpenSense: open community driven sensing of environment,": *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming*, 2010. pp. 39-42.

[2] J. A. Cohen, H. Ross Anderson, B. Ostro, D. K. Pandey, M. Krzyzanowski, N. Künzli, and, K. Smith, "The global burden of disease due to outdoor air pollution," *Journal of Toxicology and Environmental Health*, Part A, Vol. 68 (13-14), pp 1301-1307, July 2005.

[3] IEA (2016). "Energy and Air Pollution", *World Energy Outlook. International Energy Agency (IEA)*, January 10, 2017. [Online] Available: http://www.iea.org/publications/freepublications/publication/WorldEnergyOutlookSpecialReport2016EnergyandAirPollution.pdf. [Accessed February 06 2017].

[4] "ARC monographs on the evaluation of carcinogenic risks to humans". *Outdoor air pollution: International Agency for Research on Cancer* (in press), Vol. 109, 2014.

[5], J. Kim and W.J. Lee. "OpenIoT: An open service framework for the Internet of Things", *In Internet of Things (WF-IoT), IEEE World Forum*, IEEE. 2014. pp. 89-93. IEEE.

[6] N. D. Lane, E. B. Eisenmann M.. Musolesi, E. Miluzzo, and A. T. Campbell. "Urban sensing systems: opportunistic or participatory?", *In Proceedings of the 9th workshop on Mobile computing systems and applications*, ACM pp. 11-16.

[7] C. S. Lee and M. Chang. "Indoor and outdoor air quality investigation at schools in Hong Kong". *Chemosphere*, 41, no. 1, 2000. pp.109-113.

[8] F. C.Moore. "Climate change and air pollution: exploring the synergies and potential for mitigation in industrializing countries." *Sustainability* 1, no 1, 2009. pp 43-54.

[9] Moumtzidou, et al. "Towards Air Quality Estimation Using Collected Multimodal Environmental Data", I*n International Workshop on the Internet for Financial Collective Awareness and Intelligence*, Springer International Publishing, September 2016.

[10] "OpenSense II: Sensing the air we breathe". [Online] Available http://opensense.epfl.ch/wiki/index.php/OpenSense_2. [Accessed March 03 2017].

[11] "Polluscope: Participative Observatory for the Surveillance of Individual Exposure to Air Pollution and Health Effects". [Online] Available: http://polluscope.uvsq.fr. [Accessed: April 7, 2017].

[12] K. Shvachko, H. Kuang, S. Radia, and R, Chansler. "The hadoop distributed file system. In Mass storage systems and technologies (MSST)", *IEEE 26th symposium*, May 2010. pp. 1-10.

[13] S. G. Emily, T. H. Watkins, P. A. Solomon. E. D. Timothy, R. W. William, G. S.W. Heglar, D. Shelow, D. A. Hindin, V. J. Kilaru, and P. W. Preuss. "The changing paradigm of air pollution monitoring". *Environmental science & technology*, 47, no. 20, 2013.

[14] M. Walker. "Batch Vs. Real-time Data Processing". [Online] Available: http://www.datasciencecentral.com/profiles/blogs/batch-vs-real-time-data-processing. [Accessed February 02 2017].

COMING CALLS FOR PAPERS

IEEE Healthcom 2017 aims at bringing together interested parties from around the world working in the healthcare field to exchange ideas, discuss innovative and emerging solutions, and develop collaborations.

Healthcare is one of the largest industries with 5-18% of GDPs spent on health and Care globally. The healthcare budgets of the vast majority of nations continue to outgrow their GDPs. Social, demographic, economic and technological factors are the drivers for ever faster changing healthcare models. While the demographic shifts in the populations display significant socio-economic challenges they trigger more and more opportunities for innovators in the areas of sensor technology the Internet of Things, Robotics, e-health, m-Health, Cloud Computing and emerging technologies such as 5G, Big Data, SDNs, NFV, Precision and Personalized Medicine. However, the integration of innovative technology into society is associated with a lot of complexities. Social technological alignment and societal acceptance of technology requires sound solutions with regards to ethical, legal, social and security challenges.

IMPORTANT DATES:

Main Track:
Paper Submission: 1 June 2017
Acceptance Notification: 15 July 2017
Camera Ready Paper: 12 September 2017

Workshop:
Paper Submission: 25 July 2017
Acceptance Notification: 31 August 2017
Camera Ready Paper: 12 September 2017

Demo, Poster, Short Paper:
Paper Submission: 15 August 2017
Acceptance Notification: 31 August 2017
Camera Ready Paper: 12 September 2017

http://healthcom2017.ieee-healthcom.org/cfp

The International Society for Research on Internet Interventions (ISRII) is a non-profit organization of researchers, clinicians, developers, coordinators, policy makers and industry associates, among others, whose mission is to foster excellence in the development and testing of various evidence-based eHealth promotion, prevention, treatment and maintenance programs targeting behavioral and mental health to improve the health and wellbeing of individuals.

ISRII promotes the scientific study of information and communication technologies targeting behavioral, psychosocial, health and mental health outcomes. These "Internet Interventions" are broadly inclusive of existing and emerging technologies, including, but not limited to, the web, mobile and wireless devices and applications, digital gaming, virtual reality, remote sensing and robotics. ISRII members include researchers, clinicians, engineers and computer scientists, informaticists, software developers, economists, and policy experts across the public and private sectors, who are committed to fostering excellence in evidence-based eHealth interventions.
ISRII has a vibrant, enthusiastic, and highly appreciative community of members that attends and supports its conferences. Every 12 - 18 months, members come together to share findings, discuss ideas, and identify ways to collaborate.
The ISRII 9th Scientific Meeting will take place in Berlin, 12-14 October 2017. The conference's theme is *"Making e/mHealth Impactful in People's Lives"*. The meeting will have a broad scope focussing on digital technologies to improve health across the lifespan and will cover lifestyle behaviours, management of acute and chronic conditions, and mental health:

- **Methods & Applications**
- **Target Conditions**
- **Research**
- **Special populations**
- **Specific settings**

IMPORTANT DATES

- Submission deadline abstracts (required to submit a full paper later): **21-04-2017**
- Submission deadline full papers: **21-06-2017**
- Notification of acceptance: **08-08-2017**
- Camera ready version: **31-08-2017**
- Conference: **12-10-2017 – 14-10-2017**

Watch out the website for the latest news:
www.isrii-conference.com