Internet of Things Big Data Analytics: The Case of Noise Level Measurements at the Roskilde Music Festival

Tor-Morten Groenli
Mobile Technology Lab,
Dep. Technology
Kristiania University College
Oslo, Norway
tor-morten.gronli@kristiania.no

Benjamin Flesch
Dep. Digitalization
Copenhagen Business School
Copenhagen, Denmark
bf.digi@cbs.dk

Raghava Mukkamala
Dep. Digitalization
Copenhagen Business School
Copenhagen, Denmark
rrm.digi@cbs.dk

Ravi Vatrapu
Dep. Technology
Kristiania University College
Oslo, Norway
ravi.vatrapu@kristiania.no

Sindre Klavestad
Dep. Technology
Kristiania University College
Oslo, Norway
klavestad@live.no

Herman Bergner
Dep. Technology
Kristiania University College
Oslo, Norway
hermanbergner@gmail.com

***Abstract*—In this paper we demonstrate the feasibility of IoT deployment for noise level measurement to time-limited and high-intense, high-volume data, events. Through an iterative process, a prototype solution were designed and implemented in a real-time, privacy-compliant IoT sensor system under tight constraints concerning budget and development time. Our sensor system enables festival management to easily track, document and further, by applying real time big data analytics to the harvested information, have fact-full insights generated for decision making in terms of resolving noise disturbances. The whole approach was demonstrated by the use of lightweight Internet of Things architecture demonstrating how web technologies can be used throughout the technology stack in and IoT big data analytics case.**

***Keywords*-Internet of things, Big Data, Architecture, Big data analytics, IoT analytics, sound measurement, web stack.**

I. INTRODUCTION

The introduction of Internet of Things (IoT) [1] has revolutionised the access to myriads of data and paved way not only for business use cases for big data solutions, but also led to the development of the field of big data analytics. IoT embeds intelligence in the components enabling communication, exchange of information, recommendations, make decisions and provide services. The primary challenge for practitioners and researchers are still to make the most out of collected data and be able to give insights and present key findings to the primary stakeholders. From this the development has led to exploitation and exploration of big social data as well. Whilst the IoT has identified application areas such as supply chain management [2], healthcare, and workplace support, issues raised [3] by the new infrastructure include managing and making sense of the massive amount of data generated, privacy and security, as well as user acceptance of an increasingly monitored and sensor-rich world, to name but a few.

Roskilde festival depicts the largest music festival in Northern Europe, having more than 130.000 festival-goers in 2018. During the annual one-week event, a large variety of concerts both from world-renowned artists and upcoming musicians are held. As most of the festival audience stays on-site in one of the camping areas, the festival organisers are concerned with keeping up with crowd safety, visitor expectation and keeping the impact on the local society to a minimum. The organisation behind Roskilde festival depicts an unique composition, as its core staff consists of a small dedicated team of professionals who are augmented by several thousands of volunteers which need to be trained before and continuously managed during the event. In recent years, we have performed various data analytics studies at the festival, some of which have been published to the research community [4], with many of our findings reported by Danish media [5].

The purpose of this project is to record and analyse noise levels around the camping sites and camp fences. The recordings are then analysed and based on time-series data we are able to indicate whether the given noise boundaries are kept in the noise restricted areas such as in ”silent camps” or at fences close to neighbouring housing. In this paper we present the result of this project with use of Internet of Things in a big data analytics context. Furthermore our purpose is to explore how lightweight software architectures stacks can support a distributed network of devices gathering data for backend analysis and presentation.

The remainder of the paper is organised as follows. First we present the related background literature before we explain the case and the approach. Further to this, the system architecture is detailed before we present the results, discuss their impact and implications and conclude the paper with directions for future work.

II. BACKGROUND

Research into big data analytics has merged with Geographic information systems (GIS) has further extended into big data and analytics. Enabling technologies such as mobile smartphones have matured and provide real-time global positioning system (GPS) coordinates with spatial patterns that can directly feed into large-scale real-time GIS deployments [6].

Analysis and clustering of movement trajectories of individuals has been performed in various city-level case studies [7], with trajectory clustering in open-plan environments mostly focused on naval [8] and aerospace applications [9].

In its current guise, the IoT is dominated by machine-to-machine communication. The next big leap in the evolution of this communication infrastructure will be when machine to smart object communication is facilitated. The first step along this road is to create smart devices or enrich everyday objects with smart communication capabilities, laying the premises for a smart environment. Smart objects can be recognised by:

- Sensors to measure light, temperature, position etc.
- Information (data) persistence
- Communication capabilities
- Machine-to-Machine communication

IoT embeds intelligence in the components enabling communication, exchange of information, recommendations, make decisions and provide services. Although this has gained significant research and industry interest, the key challenge is to make this widespread, commonly accepted and part of the global Internet. IoT literature, expanded by the concept of fog computing [10] has increased the anchoring for such approaches.

A. Lightweight IoT Architecture

The overall vision of Internet of Things architectures as laid out by [11], where actively communicating networks creates the Internet of Things where sensors and actuators blend seamlessly with the environment. Further, Palattella et al. [12] claim that what may have previously seemed impossible given the restrictions of the Internet of Things in terms of building a standards-compliant stack may indeed become a reality. They propose a technical communication stack for an entire application, and this approach will by applying standardised approaches make this possible through open source frameworks and protocol stacks. Building on the concepts derived from Edge and Fog computing [10], we continue our efforts utilising the possibilities of a distributed, n-tiered, lightweight Internet of Things software architecture. As detailed in 1 the purpose is to push the web stack as far down into the architecture as possible, creating clear separation of concerns and enabling clear APIs between the layers [13].

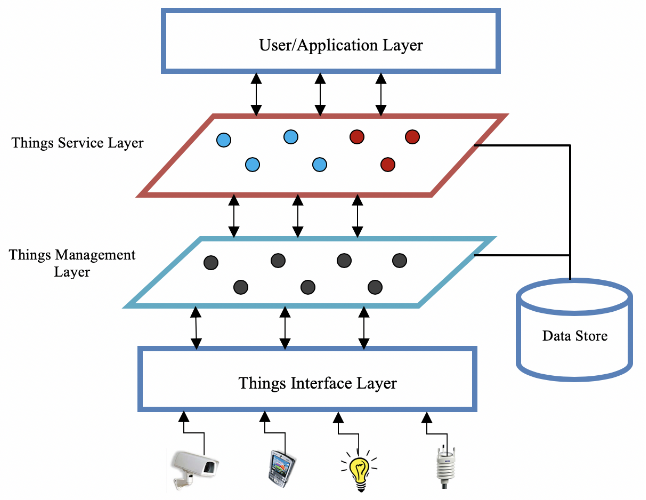

Figure 1. IoT lightweight architectural approach [13]

Antunes et al. [14] focus on the issue of data classification in IoT environment. The rationale given is that there exist billion of IoT devices which are used to generate, collect and share large volume of diversified data. However, this significantly complicates the process of managing data sources in IoT. The literature review reveals that there is no uniform way to represent, share, and understand IoT data which results in hindering the full realisation of IoT potentials. The authors propose the use of semantic information for organisation and tailoring of information. The use of such approaches, and possibly the added value from open standards such as JSON might bring forth the full potential of IoT when the devices acquire semi autonomic features and edge computing possibilities.

B. Big Data Analytics

Big Data has been traditionally characterised by the 4V's of Volume, Velocity, Variety and Veracity [15]. Big Data Analytics as a field of research and practice has been characterised by a negative distinction of data analytics that can't be accomplished by traditional mainstream analytical methods, techniques, and tools. For example, Ohlhorst [16] states that *"Big Data defines a situation in which data sets have grown to such enormous sizes that conventional information technologies can no longer effective handle either size of the data set or the scale and growth of the data set"*. Many studies have systematically reviewed the emergence and development of the principles and application soft Big Data and Big Data Analytics in specific research domains. BDA (e.g. [17], [18], [15], [19]). Current challenges in big data analytics as identified by the landmark report by the National Academy of Sciences of the United States [20] include amongst others date provenance, data security, data privacy, data sharing, and the need for real-time analysis and decision-making which are highly relevant for the use-case present in this paper.

III. Roskilde Festival 2018

Roskilde festival is located in the city of Roskilde in Denmark. It was first established in 1971, and 2018 was its 47th year of activity. Figure 2 depicts the official map of the festival area, including the live music area and camping grounds. The total area of the perimeter is 2,500,000 m2.

Figure 2. Official map of Roskilde Festival 2018

Roskilde festival 2018 lasted 8 days between Saturday 30 June and Saturday 7 July 2018 and consists of the festival area, including the iconic Orange Stage, Art Zone, Food Court and more. Headlining artists for the 47th year of Roskilde festival in 2018 were Eminem, Bruno Mars, Gorillaz, Nick Cave & The Bad Seeds, Nine Inch Nails, Massive Attack, David Byrne, Dua Lipe and Nephew.

A total of 80,000 full festival tickets and 20,000 one-day tickets were sold for the 2018 festival. The tickets were sold in 69 different countries. The biggest audiences are from Denmark, Norway, Sweden and Germany. 12.5% of the tickets were sold internationally. On top of that, more than 25.000 volunteers received free entry and camping at the festival for several days of volunteer work as part of the festival organisation. The festival is located in the small town of Roskilde, and the more than 130,000 festival-goers make Roskilde festival Denmark's fourth largest city measured by population.

IV. Design and Development

The design and development consisted of two parts. One part was the physical IoT device build in the lab to be deployed at the festival and the second part were the cloud backend with data analytics engine. Both the physical prototype and the backend application were build using iterative and incremental approach, and conducted in close cooperation with the Roskilde management representatives in order du meet the expected requirements. Further we will first detail the IoT prototype and thereafter the technical architecture of the backend system.

A. IoT Noise Level Measurement Devices

The IoT devices deployed to the field were custom build devices for sound measurement and connectivity. The devices were assembled in lab, tested and deployed. Each device consisted of a Raspberry Pi with WiFi connectivity for communication to the backend, and Arduino controlling sound level measurement. Additionally operating status and connectivity were build in using led lights, and one example device which is ready for field deployment within the festival perimeter is shown in figure 3.

Figure 3. Ardiuno-based sound measurement sensor (left background) packaged with Raspberry PI (right foreground) for IoT deployment.

The Arduino-based sound pressure level sensor is shown in figure 4. It is designed using simple components, and individually calibrated through software.

The sound pressure measurement calculation is performed within the Arduino code. This means the audio stream from the microphone attached to the Arduino is directly converted into a stream of measurement values. The output of the Arduino is a stream of measurement values in decibels, and not an audio stream. Therefore the Raspberry PI that is attached to the Arduino to provide internet connectivity never receives an audio signal, but only a stream of timestamped sound pressure level measurements. This ensures the privacy and GDPR-compliant deployment of this IoT device in a festival setting.

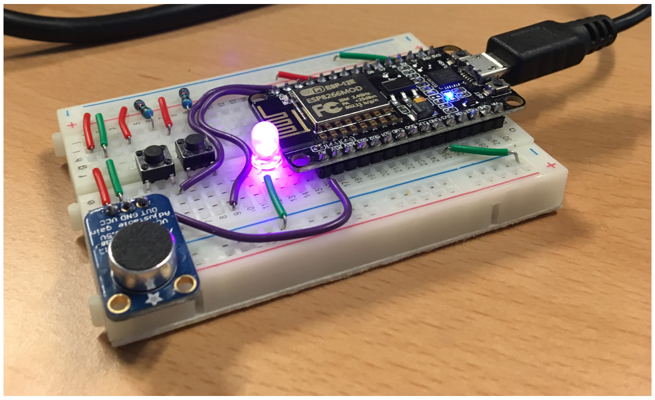

Figure 4. Arduino-based IoT sound measurement sensor

B. Application Architecture

The application architecture consists of a two clouds and a dashboard for presentation of the analysed data5. The sound measurement (of the noise levels) were done by IoT units custom build and deployed to the camp area. By utilising the concept of Fog computing / Edge computing, each deployment scenario consisted of an Arduino unit controlling the microphone and sending adapted, calculated, recordings to the local Raspberry PI gateway via WIFI. The responsibility of the gateway is to do initial filtering and aggregation of the data. By doing this, the work to be done in the backend cloud is reduced and so is bandwidth requirement for transferral of data. The initial sample rate was every two seconds, and then local aggregation was done before it was pushed to the gateway and further to the backend. Security aspects of the data is also increased as by not having the recording devices publicly accessible and the interpreted audio stream was not used ord stored any further than for converted decibel calculation.

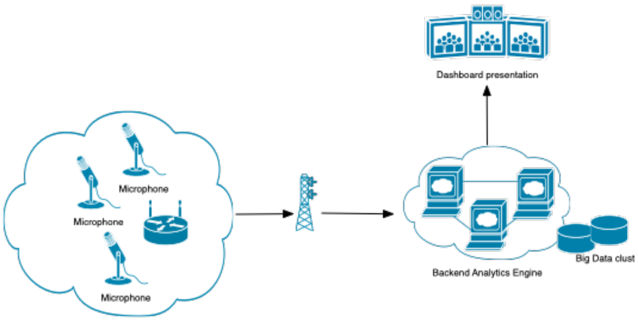

Figure 5. System Architecture

The backend cloud consists of a JavaScript Node.js implementation, with API's for gateway communication and dashboard integration. This also included the implementation of an analytics engine to process the incoming data from the different fog clouds, and ready the data for dashboard presentation. The dashboard were implemented using the Meteor framework and enabled a live streaming view, as well as daily snapshots of the actual levels. By utilising MongoDB on the backend, we were able to push JSON based data all the way from the Arduino unit to the analysis engine up to the presentation in the dashboard.

V. Result and Discussion

The results from deployed sensors were sampled on a per minute basis and presented in the form of a live stream in the dashboard, as illustrated in figure 6. Each data point consists of a timestamped sound pressure level number in decibels. Data is sent directly from the deployed IoT devices to the real-time cloud dashboard. During course of the festival, five IoT devices were deployed to the edge regions of the festival perimeter. The exact deployment locations were determined by limitations to internet access. As the first iteration of the IoT devices had no built-in means for cellular Internet access, the internet access could only be gained through the use of the festival wireless network. Further expansions would be to utilise a combined WIFI, cellular networks and narrow band GSM communication.

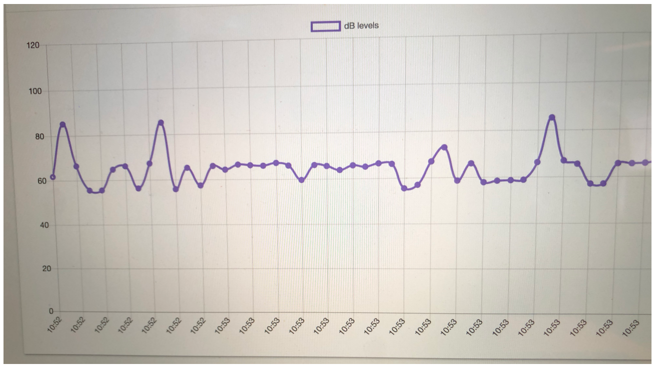

Figure 6. Live sampling from one IoT device deployed on the festival site

Further, we made use of aggregated data to show the live stream of data by the hour or by the day. Additionally the locations on the sensors were marked and visualised on a map, enabling quick identification of the relevant collection points, as showcased in figure 7.

The implications of our design choices on privacy lie within architectural choices in designing the Arduino-based noise level measurement sensor. No audio signal is stored, and only decibel measurements are communicated within the IoT device. Privacy is strengthened by this design in comparison to other design approaches [21] where a fully-fledged microphone is directly attached to the IoT device

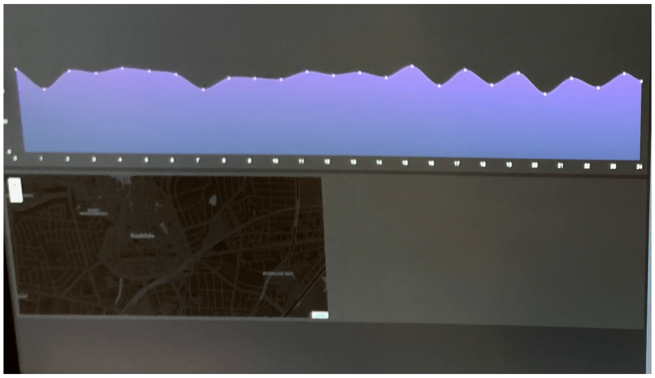

Figure 7. Live sampling from one IoT device

which has an internet connection.

In this work, we demonstrate feasibility of IoT deployment for noise level measurement to time-limited, high-intense events such as rock festivals. Through an iterative process, our research team has designed and implemented a real-time, privacy-compliant IoT sensor system under tight constraints concerning budget and development time. Our sensor system enables festival management to easily track, document and resolve noise disturbances originating from within the festival perimeters.

Typical for distributed solutions the familiar aspects of Internet of Things and edge computing include the need for multi-platform environment, device fragmentation, context-aware and timely input, low computational power and memory, energy conservation requirements, and the heterogeneity of standards. We build on our previous efforts [13] and incorporate the concept of conscious and computing capable edge nodes [11]. We further add to this, differentiating from previous research, by having the possibility for multiple edge clouds, each managing its set of sensors, which preprocess and aggregate the data before sending it to the backend cloud.

VI. Conclusion

In this work, we demonstrate feasibility of IoT deployment for noise level measurement to time-limited, high-intense events such as rock festivals. Through an iterative process, our research team has designed and implemented a real-time, privacy-compliant IoT sensor system under tight constraints concerning budget and development time. Our sensor system enables festival management to easily track, document and resolve noise disturbances originating from within the festival perimeters.

Furthermore, potential privacy concerns of audio recordings for noise level measurement are addressed by an architectural design which puts privacy first. No audio signal data is stored or even processed on the internet-connected IoT device, the Raspberry Pi. All audio signals are converted into decibel noise level measurement data right within the Arduino-based sensor, therefore a breach of privacy is not possible based on the architectural choices during design stages of the project. Therefore we enable a legally compliant deployment of the IoT devices described in this publication for use in public and semi-public scenarios.

We have in this work demonstrated how a lightweight Internet of Things architecture can be applied and how web technologies can be used throughout the technology stack. By building only on open web standards a non vendor specific approach is achieved paving path for continued open approached when combining off-the-shelf sensors, in a low cost, high performance edge cluster. When combined with real time backend analytics engine, the proposal fulfil the goal of providing real time, strategic data to base decisions upon.

For future approaches, extensive large scale implementations, further increased vendor heterogeneity and in depth threshold comparisons between the fog cloud and the back-end cloud are all worthy further pursue.

VII. Acknowledgements

This work results from our big data analytics collaboration with Roskilde festival. We thank the Roskilde festival organisers, the local authorities involved in crowd safety measures at the festival site and various volunteers from both our research group and the festival team.

References

[1] G. Kortuem, A. K. Bandara, N. Smith, M. Richards, and M. Petre, "Educating the internet-of-things generation," *Computer*, vol. 46, no. 2, pp. 53–61, 2013.

[2] S. Konomi and G. Roussos, "Ubiquitous computing in the real world: lessons learnt from large scale rfid deployments," *Personal and Ubiquitous Computing*, vol. 11, no. 7, pp. 507–521, 2007.

[3] A. Ghose, C. Bhaumik, D. Das, and A. K. Agrawal, "Mobile healthcare infrastructure for home and small clinic," in *Proceedings of the 2nd ACM international workshop on Pervasive Wireless Healthcare*. ACM, 2012, pp. 15–20.

[4] C. Zimmerman, R. Madsen, H. H. Eliassen, and R. Vatrapu, "Space vs. place: Comparing space-based movements and place-based experiences at the roskilde festival 2015," in *Proceedings of the 7th 2016 International Conference on Social Media & Society*, ser. SMSociety '16. New York, NY, USA: ACM, 2016, pp. 11:1–11:10. [Online]. Available: http://doi.acm.org/10.1145/2930971.2930983

[5] A. M. Lykkegaard, "Researchers know how you behave at roskilde festival," *CBS Wire*, Jun 2017. [Online]. Available: http://cbswire.dk/researchers-know-behave-roskilde-festival/

[6] D. Farkas, B. Hilton, J. Pick, H. Ramakrishna, A. Sarkar, and N. Shin, "A tutorial on geographic information systems: A ten-year update." *Communications of the Association for Information Systems*, vol. 38, no. 1, 2016.

[7] G. McArdle, U. Demšar, S. van der Spek, and S. McLoone, "Classifying pedestrian movement behaviour from gps trajectories using visualization and clustering," *Annals of GIS*, vol. 20, no. 2, pp. 85–98, 2014.

[8] B. Liu, E. N. de Souza, S. Matwin, and M. Sydow, "Knowledge-based clustering of ship trajectories using density-based approach," in *Big Data (Big Data), 2014 IEEE International Conference on*. IEEE, 2014, pp. 603–608.

[9] M. Gariel, A. N. Srivastava, and E. Feron, "Trajectory clustering and an application to airspace monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1511–1524, 2011.

[10] M. Aazam and E.-N. Huh, "Fog computing and smart gateway based communication for cloud of things," in *Future Internet of Things and Cloud (FiCloud), 2014 International Conference on*. IEEE, 2014, pp. 464–470.

[11] 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.

[12] M. R. Palattella, N. Accettura, X. Vilajosana, T. Watteyne, L. A. Grieco, G. Boggia, and M. Dohler, "Standardized protocol stack for the internet of (important) things," *IEEE communications surveys & tutorials*, vol. 15, no. 3, pp. 1389–1406, 2013.

[13] T.-M. Grønli, G. Ghinea, and M. Younas, "A lightweight architecture for the web-of-things," in *International Conference on Mobile Web and Information Systems*. Springer, 2013, pp. 248–259.

[14] M. Antunes, D. Gomes, and R. L. Aguiar, "Towards iot data classification through semantic features," *Future Generation Computer Systems*, vol. 86, pp. 792–798, 2018.

[15] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: from big data to big impact," *MIS quarterly*, pp. 1165–1188, 2012.

[16] F. J. Ohlhorst, *Big data analytics: turning big data into big money*. John Wiley & Sons, 2012.

[17] M. Hilbert, "Big data for development: A review of promises and challenges," *Development Policy Review*, vol. 34, no. 1, pp. 135–174, 2016.

[18] S. Akter and S. F. Wamba, "Big data analytics in e-commerce: a systematic review and agenda for future research," *Electronic Markets*, vol. 26, no. 2, pp. 173–194, 2016.

[19] A. Abbasi, S. Sarker, and R. H. Chiang, "Big data research in information systems: Toward an inclusive research agenda." *Journal of the Association for Information Systems*, vol. 17, no. 2, 2016.

[20] N. R. Council *et al.*, *Frontiers in massive data analysis*. National Academies Press, 2013.

[21] M. Willocx, J. Vossaert, and V. Naessens, "A quantitative assessment of performance in mobile app development tools," in *Mobile Services (MS), 2015 IEEE International Conference on*. IEEE, Jun. 2015, pp. 454–461. [Online]. Available: http://dx.doi.org/10.1109/MobServ.2015.68