Anomaly detection for Smart City applications over 5G Low Power Wide Area Networks

José Santos*, Philip Leroux*, Tim Wauters*, Bruno Volckaert* and Filip De Turck*
*Ghent University - imec, IDLab, Department of Information Technology
Technologiepark-Zwijnaarde 15, 9052 Gent, Belgium
Email: josepedro.pereiradossantos@UGent.be

***Abstract*— In recent years, the Internet of Things (IoT) has introduced a whole new set of challenges and opportunities in Telecommunications. Traffic over wireless networks has been increasing exponentially since many sensors and everyday devices are being connected. Current networks must therefore adapt to and cope with the specific requirements introduced by IoT. One fundamental need of the next generation networked systems is to monitor IoT applications, especially those dealing with personal health monitoring or emergency response services, which have stringent latency requirements when dealing with malfunctions or unusual events. Traditional anomaly detection approaches are not suitable for delay-sensitive IoT applications since these approaches are significantly impacted by latency. With the advent of 5G networks and by exploiting the advantages of new paradigms, such as Software-Defined Networking (SDN), Network Function Virtualization (NFV) and edge computing, scalable, low-latency anomaly detection becomes feasible. In this paper, an anomaly detection solution for Smart City applications is presented, focusing on low-power Fog Computing solutions and evaluated within the scope of Antwerp's City of Things testbed. Based on a collected large dataset, the most appropriate Low Power Wide Area Network (LPWAN) technologies for our Smart City use case are investigated.**

***Index Terms*—Anomaly Detection, Smart Cities, IoT, 5G, LPWAN, Fog Computing**

I. Introduction

In recent years, with the advent of the Internet of Things (IoT), the concept of Smart Cities has become even more popular [1]. IoT will transform a wide range of services in different domains of urban life, by creating intelligent smart grids, improving public transportation and developing car parking and personal health monitoring applications. In the future network generation, information will be transmitted from different types of devices, over heterogeneous wireless networks with even higher data rates, lower latencies and lower power consumption [2]. Therefore, it will be necessary to adapt existing network architectures to future needs and develop new autonomous management functionalities to help meet the demanding requirements of future 5G use cases. In fact, 5G technologies promise very high carrier frequencies with massive bandwidths, extreme base station densities, an unprecedented large numbers of antennas and new functionalities, such as device-to-device communication (D2D) and Fog Computing [3], [4]. In Fig. 1, a 5G network architecture is presented in a Smart City context. 5G technologies aim to tackle the new business opportunities created by the stringent requirements of IoT applications. One of the main challenges is how to efficiently handle with the gathering and processing of all data coming from the enormous amount of IoT sensors that will be connected to the network in the next years [5].

978-1-5386-3416-5/18/$31.00 © 2018 IEEE

The Fog Computing paradigm, which places cloud resources close to the IoT sensors, extends the Cloud Computing paradigm to deal with the eminent growth of connected devices [6]. Nevertheless, Fog Computing is still in its early stages and needs more time to evolve. One of the remaining challenges is how to provide proper resource allocation, since IoT applications and services can be placed in a highly congested area, which would result in a higher latency [7]. Furthermore, current IoT sensors and gateways lack in terms of processing power, battery, memory and storage capacity [8], [9]. IoT applications will be so diverse that they will have different sets of communication requirements. For instance, on one hand, a delay-sensitive IoT application may require very low latencies, meaning this IoT application must be allocated on fog resources close to the sensor enabling the control of time-sensitive network functionalities close to the device [10]. On the other hand, if this requirement is less important, the IoT application could be placed far from the IoT sensor in a central location in order to reduce the number of active fog resources on the network and therefore minimize the total energy consumption in the fog domain. Additionally, it is also important to detect malfunctions and abnormal events in the network. By identifying unusual events, malfunctions in IoT sensors can be detected and transmissions of incorrect information can be avoided, which can improve the overall Quality of Service (QoS) of the IoT application, especially in terms of reliability [9]. Detecting unexpected patterns in the data traffic is known as anomaly detection [11]. Recently, anomaly detection has attracted the attention of the research community in multiple areas, such as intrusion detection, health monitoring, preventive maintenance and fault detection [12]. In this paper, an anomaly detection approach for IoT applications in 5G Smart Cities based on the advantages of Fog Computing architectures is presented. The proposed architecture has been designed for Antwerp's City of Things testbed [13] and validated for Smart City use cases, in particular for an Air Quality monitoring application. Finally, multiple Low Power Wide Area Network (LPWAN) technologies have been considered for our use case scenario. Our evaluation results

identify the most adequate LPWAN technologies as wireless communication enablers for the considered IoT application.

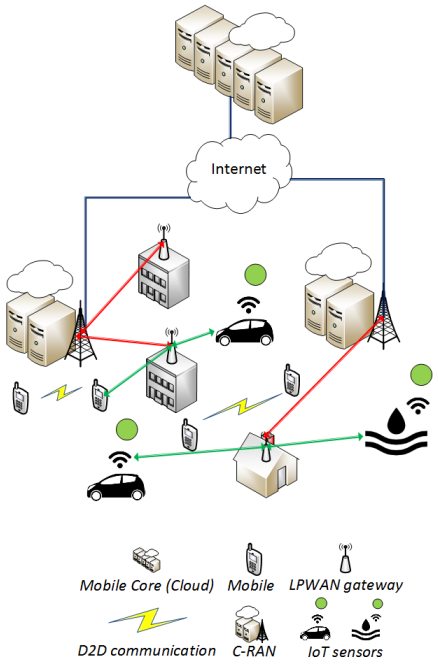


Fig. 1: High-level view of the considered 5G network architecture.

The remainder of this paper is organized as follows. In the next section, related work is discussed. In section III, the anomaly detection approach for smart city environments as well as the LPWAN network dimensioning is presented. Then, in Section IV, the scenario and datasets used in the evaluation are presented, followed by the evaluation results in section V. Finally, conclusions are presented in section VI.

II. Related Work

Recently, research studies have been carried out in order to deal with anomaly detection in IoT, Smart City and Industry 4.0 scenarios. In [14], a real-time Intrusion Detection System (IDS) for IoT has been presented. The proposed solution is a novel IDS with an integrated mini-firewall for 6LoWPAN networks in order to detect malicious nodes. Moreover, in [11], an anomaly detection scheme based on sensor data has been proposed to deal with unexpected behaviors in turbomachines in the Petroleum Industry. Furthermore, in [15], a temporal clustering and anomaly detection method has been presented for a car parking IoT application in order to detect unusual events. In [16], a supervised statistical-based anomaly detection method for Smart Grid data has been proposed.

In recent years, research projects have also been focusing on reliable and secure IoT for Smart Cities. In the SOCIOTAL project [17], an anomaly detection method based on hyperellipsoidal models has been used to identify unusual patterns in environmental data collected from IoT sensors [18]. However, a traditional cloud solution has been deployed instead of a Fog Computing approach. Additionally, in [9], a Fog Computing anomaly detection approach for IoT using a hyperellipsoidal clustering algorithm has been proposed to significantly reduce latency and energy consumption in the network when compared to distributed and centralized architectures. Nevertheless, their work is based on simulation studies, while our approach is based on an actual deployment within the scope of Antwerp's City of Things testbed. Finally, in the CityPulse project [19], [20], a complete set of real-time data analytics tools have been presented, such as data aggregation, event detection and decision support.

In summary, in this paper, a Fog-based anomaly detection approach is proposed. Our work takes into account not only the advantages of Fog Computing architectures, which are suitable for IoT applications in Smart City scenarios, but also characteristics stemming from LPWAN technologies, since low power technologies have gained tremendous emphasis due to the enormous growth of connected devices. All Anomaly detection approaches cited are only focused on cloud and data management aspects and no considerations are included about wireless networks. The proposed approach has been implemented on our City of Things platform and evaluated with Unsupervised Clustering and Outlier detection algorithms for our Air Quality monitoring application. Furthermore, the most popular LPWAN technologies today have been assessed based on the requirements of our application.

III. Anomaly Detection in Smart Cities

This section presents the proposed low-latency anomaly detection approach for Smart Cities on Fog Computing resources interconnected by LPWAN networks.

A. Anomaly Detection Principles

Anomaly detection or outlier detection is known as the process of detecting unexpected behavior or abnormal patterns in datasets. In the past, anomaly detection was mainly used to remove the outliers from a dataset, which is called data cleansing. However, in recent years, anomaly detection has attracted the attention of the research community because researchers began to get interested in knowing more about the anomalies themselves, since they are usually associated with potentially reoccurring events [21]. There are three main categories of anomaly detection which are shown in Fig. 2.

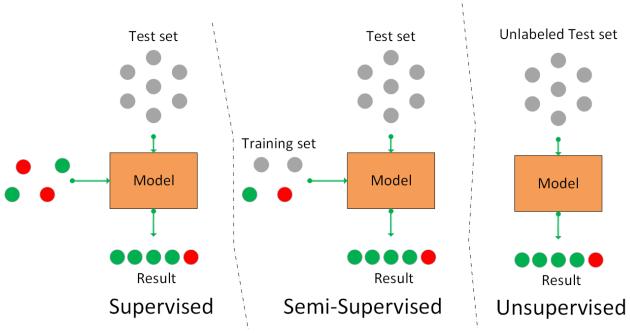


Fig. 2: Anomaly Detection Categories: Supervised, Semi-supervised and Unsupervised.

1) Supervised: In supervised anomaly detection methods, a fully labeled training dataset is used, i.e., each sample is considered as normal or abnormal. This category is only used for specific applications where anomalies are known beforehand.

2) Semi-Supervised: In semi-supervised anomaly detection techniques, a training dataset consists of labeled and unlabeled samples. Usually, a small amount of labeled samples is used.

3) Unsupervised: In unsupervised anomaly detection algorithms, there is no training dataset since nothing is known about the samples in advance. Therefore, these methods usually give an estimation of what a normal sample and what an abnormal one is. The approach presented in this paper makes use of unsupervised techniques since our goal was to see if the selected algorithms could learn the distribution of the data samples without knowing anything beforehand as in most anomaly detection use cases. Our objective was that the algorithms themselves could discover and present interesting behaviors in the datasets. Moreover, labeling datasets for anomaly detection is not an easy task.

Many categories of unsupervised anomaly detection algorithms exist, of which the most popular are listed in Table I. Many of these algorithms are available in Scikit-Learn [22], a powerful machine learning library written in Python, which has been used to implement the anomaly detection evaluations.

TABLE I: Unsupervised Anomaly Detection Algorithms [23]

Statistical-based	Univariate and Multivariate Gaussian distribution, Grubbs' test, Likelihood approach
Proximity-based	K Neighbors
Clustering-based	K-Means, MiniBatchKMeans, Birch
Density-based	Local Outlier Factor (LOF)
One-class support vector machines	One Class SVM, Gaussian envelope (Robust Covariance)
Ensemble-based	Isolation Forrest (IF)

B. Fog-based Anomaly Detection Approach

To deal with the growing amount of connected devices in the network, the Fog Computing paradigm has been introduced to place computational resources on the edges of the network in order to deal with the stringent requirements introduced by IoT use cases, such as low latency and high mobility. Centralized solutions are not suitable for most future IoT applications, since most of them will require real-time communication and generate an enormous volume of data to be transported in the network, which makes it impossible for centralized solutions to comply with these requirements. The IoT sensors communicate with wireless gateways, which are linked with the fog resource layer, managing a set of computational resources. These fog resources can communicate with the cloud layer, which is the top level management entity. Each service or IoT application must be allocated to and instantiated on a given set of computational resources. For instance, for a delay-sensitive IoT application, service allocation must be performed on a fog resource as close as possible to the IoT sensor that runs the client application allowing real-time processing and data analytics at the edges of the network in order to enable the control of time-sensitive network functionalities close to the IoT sensor.

In a traditional centralized approach, all IoT sensors send their data samples to the cloud layer. Then, the anomaly detection algorithms are executed. This approach implies a high bandwidth cost, because all data samples need to be transmitted from the IoT sensors to the cloud layer. Our approach presented in Fig. 3 is based on the advantages of Fog Computing architectures, i.e., anomaly detection operations are performed on fog resources. Every IoT sensor sends its data samples to one of the fog resources. Then, anomaly detection operations are performed in a distributed way. After completion of the anomaly detection tasks, fog resources may send alerts to the cloud layer and to the IoT sensors if unusual events are already detected on the data. This way, faster response times can be achieved if any abnormal behavior is discovered. Moreover, fog resources can send the outcomes of the anomaly operations to the cloud layer to combine results from the different fog resources in order to have a broader view of the behavior of the network. Afterwards, the cloud layer could perform global anomaly detection operations and present the outcomes in a central dashboard in a control room. Then, alerts can be sent to fog resources and IoT sensors in case abnormal patterns or inconsistent events were not detected.

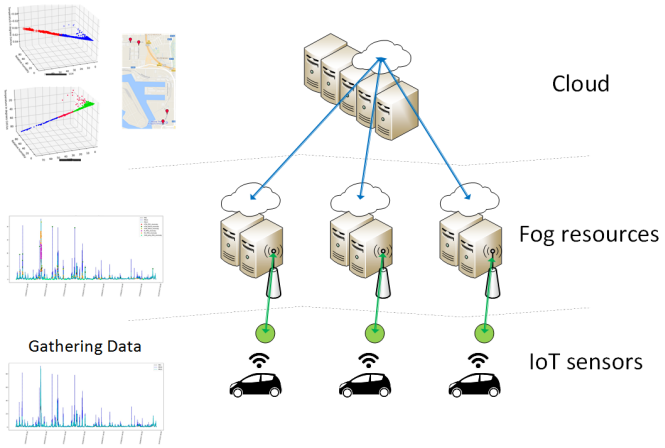


Fig. 3: High-level view of the Fog-based anomaly detection approach.

C. LPWAN Dimensioning

Nowadays, low power wireless technologies have gained tremendous emphasis due to the massive growth of connected devices in the network. The need for connecting simple IoT devices, such as sensors and actuators, is increasing rapidly. In Table. II, the most popular LPWAN technologies and the main differences between them are shown. To select a suitable LPWAN technology for a specific IoT application, an analysis of its requirements in terms of specific parameters such as communication range, upload and download data rate, frequency bands, and latency is needed. In this paper,

TABLE II: Comparison between different LPWAN technologies for IoT applications [24], [25], [26], [27]

LPWAN Technology	LoRaWAN	Sigfox	LTE-M	DASH7	IEEE 802.11ah	NB-IoT	Ingenu RPMA
Range urban	2-5 km	3-10 km	2-5 km	5 km	1 km	2-5 km	1-3 km
Range rural	15 km	30-50 km	-	5 km	1 km	-	25-50 km
Data rate	50 kbps	300bps	1 Mbps	166.67 kbps	346.66 Mbps	250 kbps	634 kbps (uplink) 156 kbps (downlink)
Bi-directional	Yes	Limited	Yes	Yes	Yes	Yes	Yes
Freq. band	Unlicensed	Unlicensed	Licensed	Unlicensed	Unlicensed	Licensed	Unlicensed
Power efficiency	Very high	Very high	Medium	Medium	High	Medium	Medium
Security	Medium	Low	High	Medium	Low (In development)	High	High
Mobility	Yes	Limited	Yes	Yes	Yes	Yes	Yes
Proprietary	No	Yes	No	No	No	No	Yes

multiple LPWAN technologies have been evaluated based on the requirements of our Air Quality monitoring use case presented in Section IV-A.

Variables used in our LPWAN dimensioning are shown in Table III. In Fog Computing architectures, fog resources are usually located within one hop from the IoT sensors. The variable C is used to indicate the communication range in kilometers between a fog resource and an IoT sensor. Two variables, U and D, are used to indicate the upload and the download data rate, respectively. Then, the total number of bits to be transmitted is given by N. This way, the upload and the download transmission time of a packet can be expressed as shown in (1) and in (2), respectively.

$$T(upload) = \frac{N}{U} \tag{1}$$

$$T(download) = \frac{N}{D} \tag{2}$$

Moreover, by using the communication range C and the propagation speed P for wireless communications, which is the speed of light (3×10^8), the propagation time is given by (3).

$$P = \frac{C}{3 \times 10^8} \tag{3}$$

Finally, the total packet delivery time L is given by (4).

$$L = T + P \tag{4}$$

TABLE III: Variables of the LPWAN dimensioning

Symbol	Description
C	Communication range in kms.
U	Upload Data Rate in kbps.
D	Download Data Rate in kbps.
N	Number of bits to be transmitted.
T	Transmission time of a packet.
P	Propagation time of a packet.
L	Packet Delivery time.

IV. Evaluation Scenario

In this section, the evaluation scenario is introduced. Then, the datasets are presented. Finally, the evaluation setup is described.

A. Use Case - Air Quality Monitoring Application

The evaluation scenario is based on a use case within the scope of Antwerp's City of Things testbed. The goal of our Air Quality monitoring application is to detect high amounts of organic compounds in the atmosphere and then alert citizens of air pollution in real-time. As an initial proof of concept, Air Quality sensors have been integrated in collaboration with the Belgian postal services Bpost [13]. For daily mail delivery, Bpost has cars driving around in the city of Antwerp. Therefore, within the City of Things project, a set of Air Quality sensors have been mounted on the roofs of Bpost's delivery cars as shown in Fig. 4. These sensors send measures of typical gases and climate data such as temperature and humidity, which are then annotated with GPS locations. Moreover, these sensors allow gathering real-time Air Quality information with broad city coverage, since each car is continuously driving around in the city. Currently, the set of Air Quality sensors can communicate via three different LPWAN technologies: LoRaWAN, SigFox and DASH7 [13].

B. Datasets

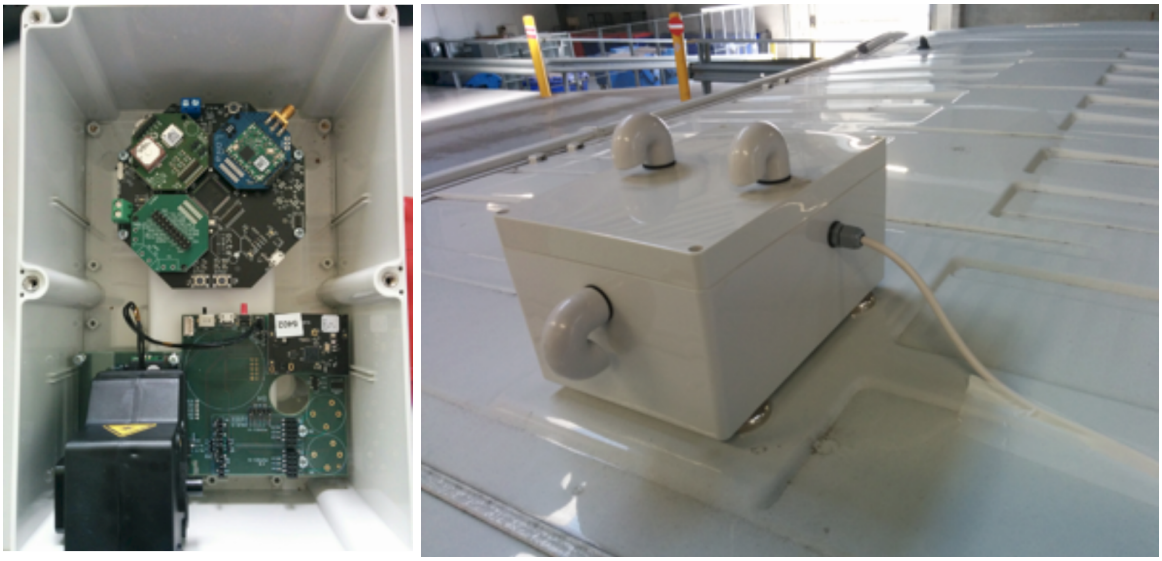
A summary of the characteristics of the datasets gathered for the evaluation is shown in Table IV. The two datasets come from two different Bpost cars and consist of particle matter indicators (PM1, PM2.5 and PM10) that are annotated with a GPS location. The datasets have been collected by our research group between 2017-05-09 and 2017-06-29. The particle matter indicators are shown in Fig. 5a and in Fig. 5b for Bpost car 1 and Bpost car 2, respectively.

TABLE IV: Evaluation datasets

Dataset Name	No. of records	Description
Bpost 1	70636	Particle matter indicators (PM1, PM2.5, PM10) and GPS locations from Bpost car 1 between 2017-05-09 and 2017-06-29
Bpost 2	70640	Particle matter indicators (PM1, PM2.5, PM10) and GPS locations from Bpost car 2 between 2017-05-09 and 2017-06-29

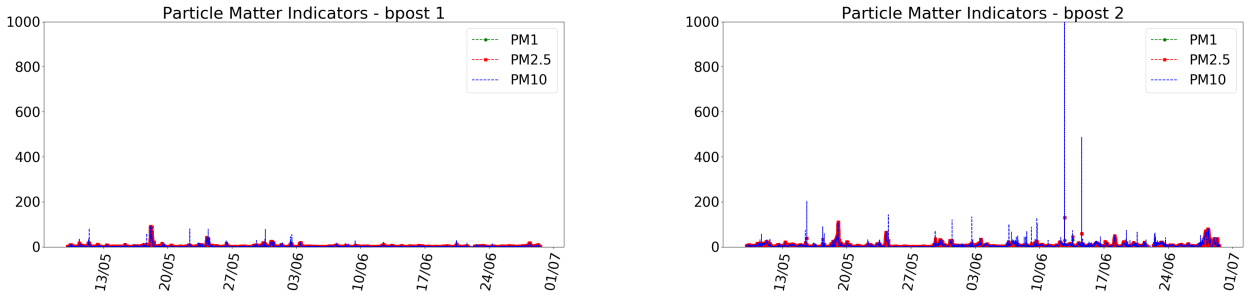
C. Selected Algorithms

As previously mentioned, the anomaly detection evaluations have been implemented in Python using Scikit-Learn. Unsu-



(a) Inside view of the multi-radio sensor. (b) Air Quality sensor mounted on a Bpost car.

Fig. 4: As part of the Antwerp's City of Things testbed, multi-radio Air Quality sensors have been mounted on cars of the Belgian postal service.



(a) Particle Matter Indicators (PM1, PM2.5, PM10) - Bpost car 1 (b) Particle Matter Indicators (PM1, PM2.5, PM10) - Bpost car 2

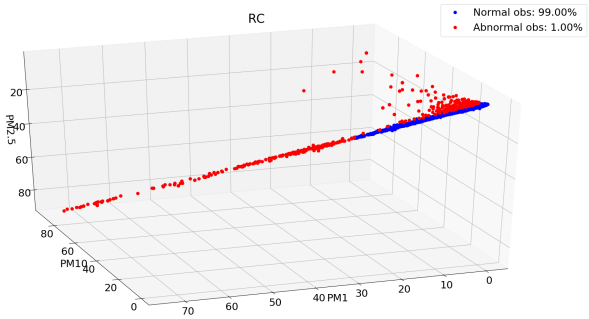
Fig. 5: Particle Matter Indicators (PM1, PM2.5, PM10 in ppm) for two Bpost cars.

pervised Clustering and Outlier detection algorithms have been assessed by using the two datasets presented in Section IV-B. Clustering allows the detection of patterns in unlabeled data with many dimensions while Outlier detection is related to the identification of unusual data samples when compared to the rest of the dataset. Regarding Clustering, the Birch algorithm has been evaluated while for Outlier detection Robust Covariance (RC) has been assessed. Birch and RC outcomes have been compared in order to find patterns or unusual events in the datasets. Furthermore, the results have been compared with the correspondent GPS locations to know exactly where in the city of Antwerp each sample has been measured.

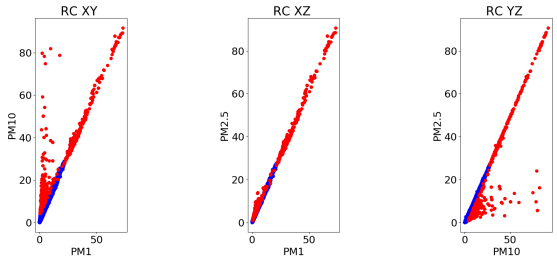
V. Evaluation Results

A. Clustering and Outlier Detection

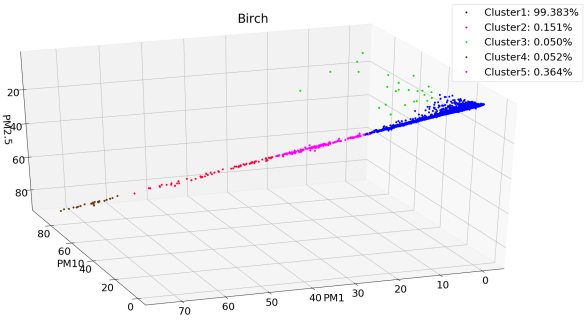
In Fig. 6a and in Fig. 6b, the outcomes of the RC outlier detection algorithm for the three dimensions regarding particle matter indicators for the Bpost car 1 are shown with a contamination of 1.0% indicating that the RC algorithm intends to find the 1.0% of samples which can be considered as abnormal. Moreover, in Fig. 6c and in Fig. 6d, the results obtained for the Bpost car 1 by using the Birch clustering algorithm for 5 clusters are illustrated. Regarding outlier detection, values of PM1, PM2.5 and PM10 above 30 ppm collected by Bpost car 1 are marked as outliers meaning that these values can be considered as unusual. Regarding clustering, there are clear similarities between these results and the outcomes obtained by the RC outlier detection algorithm. The first cluster is composed of almost 99.4% of the total data samples indicating that this cluster can be considered as the normal region of values for the three particle matter indicators collected by Bpost car 1. Moreover, the second and the fifth cluster, consist of 0.5% of the data samples which are also detected as outliers by the RC algorithm. These clusters can be considered as unusual regions. Finally, the third and the fourth cluster, are



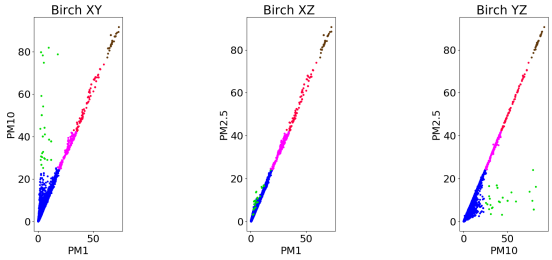
(a) RC with 1.0% - 3D perspective - Bpost car 1



(b) RC with 1.0% - 3D planes - Bpost car 1

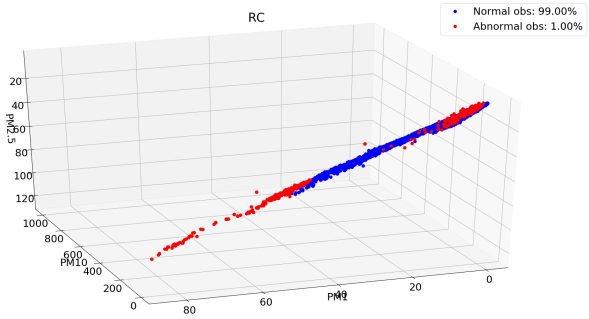


(c) Birch: 5 clusters - 3D perspective - Bpost car 1

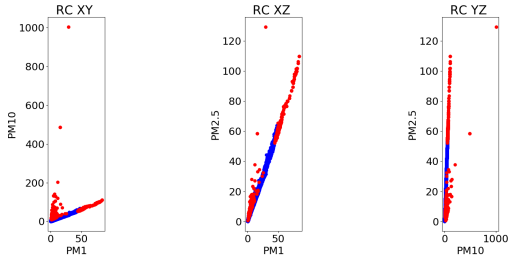


(d) Birch: 5 clusters - 3D planes - Bpost car 1

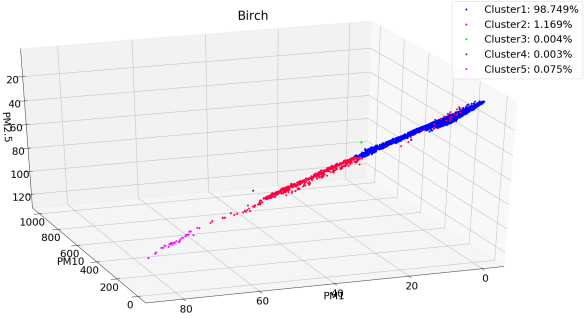
Fig. 6: Robust Covariance Outlier detection with a contamination of 1.0% (blue color: normal samples, red color: abnormal samples) and Birch Clustering results with 5 clusters (blue, red, green, brown, pink) for Particle Matter Indicators (PM1, PM2.5, PM10) - Bpost car 1.



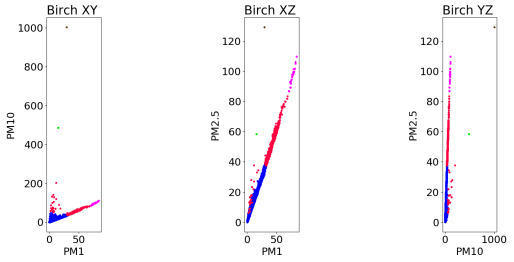
(a) RC with 1.0% - 3D perspective - Bpost car 2



(b) RC with 1.0% - 3D planes - Bpost car 2



(c) Birch: 5 clusters - 3D perspective - Bpost car 2



(d) Birch: 5 clusters - 3D planes - Bpost car 2

Fig. 7: Robust Covariance Outlier detection with a contamination of 1.0% (blue color: normal samples, red color: abnormal samples) and Birch Clustering results with 5 clusters (blue, red, green, brown, pink) for Particle Matter Indicators (PM1, PM2.5, PM10) - Bpost car 2.

composed of 0.1% of data samples which are also detected as outliers by the RC algorithm. These clusters can therefore be considered as very unusual regions.

In Fig. 7a and in Fig. 7b, the outcomes of the RC algorithm with a contamination of 1.0% for the three dimensions regarding particle matter indicators for the Bpost car 2 are shown. In Fig. 7c and in Fig. 7d, the results obtained for the Bpost car 2 by using the birch clustering algorithm for 5 clusters are illustrated. Regarding outlier detection, values of PM1 above 60 ppm, PM2.5 above 70 ppm and values of PM10 above 150 ppm collected by Bpost car 2 are marked as outliers. This way, these values can be considered as unusual data samples. Regarding clustering, there are clear similarities between these results and the outcomes obtained by the RC outlier detection algorithm, as shown for the Bpost car 1.

B. GPS Locations of outliers

Outliers must be further analyzed by application experts in order to extract more information from them. In our evaluation, the outliers have been compared with the GPS locations available in the datasets. In Fig. 8, the GPS locations are shown where PM10 values above 75 ppm and PM2.5 values above 30 ppm have been collected by the Bpost cars, which have been considered as unusual data samples by the RC outlier detection algorithm.

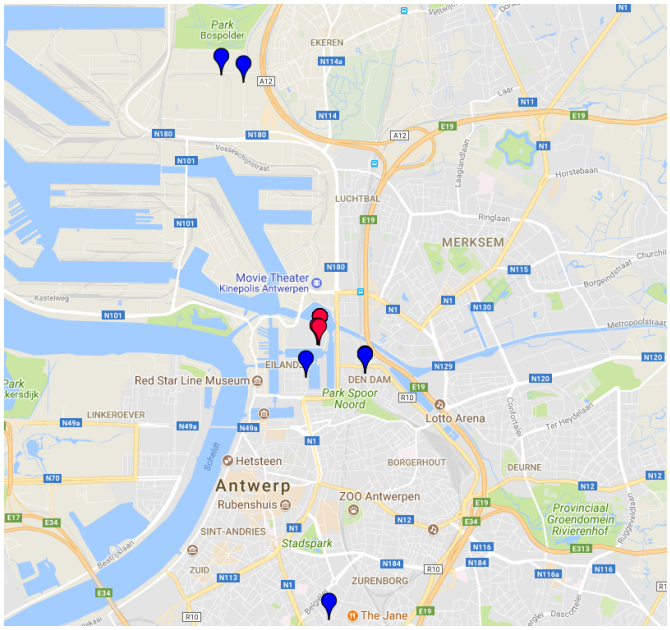


Fig. 8: GPS locations (Bpost car 1 - red / Bpost car 2 - blue) considered as outliers by the RC algorithm.

Regarding Bpost car 1 measurements, all the unusual values have been collected in the warehouse where usually the Bpost cars stay at night. In fact, all these values have been collected on a single night between 2:54 am until 6:33 am on 5/18/2017. These high values of PM10 and PM2.5 can be related to dust and organic compounds, which were inside the warehouse at the time of the measurements. On the other hand, the unusual values measured by Bpost car 2 have been collected across the city of Antwerp. These high values of PM10 and PM2.5 can be explained by high traffic volumes in the city at these locations at the time of the measurements. This way, by conducting anomaly detection operations in fog resources, timely alerts can be transmitted to IoT sensors and to the cloud layer indicating that high values of particle matter indicators have been measured. In doing so, citizens can be alerted of high air pollution levels in real-time.

C. LPWAN Dimensioning Analysis

Considering that for our use case, each upload message is composed of a String of 12 chars (GPS Location - geohash) equal to 12 bytes, a 32 bit integer (timestamp) equal to 4 bytes and 3 floating point 64 bit numbers (particle matter indicators) equal to 24 bytes, the total number of payload bytes to be transmitted per minute from the IoT sensor to the fog resource is 40 bytes. On the other hand, each download message to be transmitted from the fog resource to the IoT sensor in case of unusual behavior or malfunction is composed of a String of 12 chars (GPS Location - geohash) equal to 12 bytes and a byte defined by 3 alarm bits and 5 bits for 32 types of predefined messages. Furthermore, each message has a header for which the size depends on the LPWAN technology itself. In our evaluation, a general 13 byte header has been considered in each message as in Sigfox and in LoRaWAN technologies. Therefore, each upload message is transmitted with at least 53 bytes which is equal to 424 bits and each download message with 26 bytes which is equal to 208 bits. Moreover, considering that the area of Antwerp is equal to 204.5 km^2, an estimation of the minimum number of gateways required for each LPWAN technology to cover the entire area of the city has been performed. Based on these assumptions, the LPWAN technologies presented in Table II have been evaluated. The comparison is presented in Table V and in Table VI, a list of pros and cons for the multiple LPWAN technologies is shown. Based on these results and because of our application requirements, Sigfox technology is unfeasible to provide wireless communication, since a single upload message takes more than a second to be transmitted and due to duty cycle regulations, real-time communication is not possible, because sending an upload message every minute implies going against the fairness rules of duty cycle. Moreover, the download capabilities are very limited in Sigfox technology. On the other hand, Ingenu RPMA is not considered as an adequate solution because it operates in the crowded 2.4 GHz band. Nowadays, low frequencies are being considered as optimal to provide wireless communication for IoT solutions. Besides, Ingenu RPMA requires high processing power, which translates into a higher energy consumption.

Regarding licensed LPWANs, LTE-M is the optimal solution to deploy our use case, because it has a higher data rate than NB-IoT making LTE-M more suitable for our application since real-time communication is needed for our scenario. On the other hand, regarding unlicensed LPWANs, IEEE 802.11ah and DASH7 are the most adequate solutions to provide wire-

TABLE V: Comparison between the different LPWAN technologies based on the requirements of the Air Quality application

LPWAN Technology	LoRaWAN	Sigfox	LTE-M	DASH7	IEEE 802.11ah	NB-IoT	Ingenu RPMA
C	5 km	10 km	5 km	5 km	1 km	5 km	3 km
U/D	50 kbps	300bps	1 Mbps	166.67 kbps	346.66 Mbps	250 kbps	634/156 kbps
P	16.67ms	33.33ms	16.67ms	16.67ms	3.33ms	16.67ms	10ms
N (upload)	At least 53 bytes						
T (upload)	8.480ms	1.413s	0.424ms	2.543ms	1.221μs	1.696ms	0.669ms
L (upload)	25.15ms	1.45s	17.09ms	19.22ms	3.331ms	18.37ms	10.67ms
N (download)	At least 32 bytes						
T (download)	4.16ms	0.69s	0.208ms	1.24ms	0.60μs	0.83ms	1.33ms
L (download)	20.83ms	0.72s	16.88ms	17.91ms	3.331ms	17.5ms	11.33ms
Area of the City of Antwerp	204.5 km^2						
Minimum Number of Gateways	5	2	5	5	117	5	15

TABLE VI: List of Pros and Cons of the different LPWAN technologies for the Air Quality application

LPWAN Technology	PROS	CONS
LoRaWAN	Security;	Limited downlink capability;
Sigfox	None;	Duty cycle regulations (transmit time of 36s per 1 hour): impossible to transmit a message every minute for our application; Proprietary protocol; Limited security; Limited downlink;
LTE-M	High data rate; Security;	Under development;
DASH7	High data rate when compared with similar LPWANs;	Open Source Solution; Lack of deployments;
IEEE 802.11ah	Lower transmission time; High data rate;	High number of gateways needed (low range when compared with other LPWANs); Under development;
NB-IoT	High data rate when compared with other LPWANs; Security;	Under development;
Ingenu RPMA	High uplink data rate; High coverage and robustness;	Lower range when compared with other LPWANs; Operates in the crowded 2.4Ghz band; High processing power;

less communication between the devices. LoRaWAN is not considered as an appropriate solution, because the downlink capacities are very limited. Moreover, LoRaWAN has the lower data rate and the correspondent highest transmission time when compared with IEEE 802.11ah and DASH7. In fact, although IEEE 802.11ah is currently under development, it is one the most promising LPWAN technologies with a very high data rate. However, IEEE 802.11ah deployment will require a very large number of gateways to cover the entire city of Antwerp due to the low communication range. On the other hand, DASH7 supports high data rates when compared to similar LPWAN technologies and it is already deployed in the City of Things testbed. Both technologies meet our application demands, which make them appropriate solutions to provide wireless communication for our use case in the unlicensed spectrum.

VI. Conclusions

In recent years, the need for management functionalities in Smart Cities is increasing due to the deployment of IoT use cases. Fog Computing provides an efficient manner of dealing with stringent requirements introduced by IoT use cases. It is important to detect malfunctions and abnormal events in IoT devices to provide a secure and reliable communication. Therefore, in this paper, a low-latency Fog-based anomaly detection approach has been presented to identify unusual events or abnormal patterns in IoT scenarios. Our approach has been evaluated for a Smart City use case within the scope of City of Things testbed. Obtained results show that both Birch clustering and RC outlier anomaly detection mechanisms can be performed by fog resources close to IoT sensors and, this way, send timely alerts in case unusual events are detected. Moreover, for multiple criteria, LPWAN technologies have been evaluated for our Air Quality application, leading to a suitable set of LPWAN technologies, IEEE 802.11ah, DASH7 and LTE-M, that can be used as wireless communication enablers for our Smart City use case. As future work, the selected LPWAN technologies will be deployed and techno-economical studies will be performed.

Acknowledgment

This research was performed as part of the "City of Things" project (Antwerp, Belgium) funded by imec and in the "Intelligent DEnse And Longe range IoT networks (IDEAL-IoT)" project under Grant Agreement #S004017N, from the fund for Scientific Research-Flanders (FWO-V).

References

[1] V. Albino, U. Berardi, and R. M. Dangelico, "Smart cities: Definitions, dimensions, performance, and initiatives," *Journal of Urban Technology*, vol. 22, no. 1, pp. 3–21, 2015.

[2] T. Han, X. Ge, L. Wang, K. S. Kwak, Y. Han, and X. Liu, "5g converged cell-less communications in smart cities," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 44–50, 2017.

[3] A. Gupta and R. K. Jha, "A survey of 5g network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[5] (2017) Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf

[6] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[7] M. Taneja and A. Davy, "Resource aware placement of iot application modules in fog-cloud computing paradigm," in *Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on*. IEEE, 2017, pp. 1222–1228.

[8] H. I. Kobo, A. M. Abu-Mahfouz, and G. P. Hancke, "A survey on software-defined wireless sensor networks: Challenges and design requirements," *IEEE Access*, vol. 5, pp. 1872–1899, 2017.

[9] L. Lyu, J. Jin, S. Rajasegarar, X. He, and M. Palaniswami, "Fog-empowered anomaly detection in internet of things using hyperellipsoidal clustering," *IEEE Internet of Things Journal*, 2017.

[10] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," *IEEE Network*, vol. 30, no. 4, pp. 46–53, 2016.

[11] L. Martí, N. Sanchez-Pi, J. M. Molina, and A. C. B. Garcia, "Anomaly detection based on sensor data in petroleum industry applications," *Sensors*, vol. 15, no. 2, pp. 2774–2797, 2015.

[12] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *Ieee communications surveys & tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[13] S. Latre, P. Leroux, T. Coenen, B. Braem, P. Ballon, and P. Demeester, "City of things: An integrated and multi-technology testbed for iot smart city experiments," in *Smart Cities Conference (ISC2), 2016 IEEE International*. IEEE, 2016, pp. 1–8.

[14] S. Raza, L. Wallgren, and T. Voigt, "Svelte: Real-time intrusion detection in the internet of things," *Ad hoc networks*, vol. 11, no. 8, pp. 2661–2674, 2013.

[15] Y. Zheng, S. Rajasegarar, C. Leckie, and M. Palaniswami, "Smart car parking: temporal clustering and anomaly detection in urban car parking," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on*. IEEE, 2014, pp. 1–6.

[16] X. Liu and P. S. Nielsen, "Regression-based online anomaly detection for smart grid data," *arXiv preprint arXiv:1606.05781*, 2016.

[17] (2017) SOCIOTAL project. An EU FP7 funded STREP project addressing the objective FP7-ICT-2013.1.4 "A reliable, smart and secure Internet of Things for Smart Cities". [Online]. Available: http://www.sociotal.eu

[18] P. Rathore, A. S. Rao, S. Rajasegarar, E. Vanz, J. Gubbi, and M. Palaniswami, "Real-time urban microclimate analysis using internet of things," *IEEE Internet of Things Journal*, 2017.

[19] (2017) CityPulse: Real-Time IoT Stream Processing and Large-scale Data Analytics for Smart City Applications. [Online]. Available: http://www.ict-citypulse.eu

[20] D. Puiu, P. Barnaghi, R. Toenjes, D. Kümper, M. I. Ali, A. Mileo, J. X. Parreira, M. Fischer, S. Kolozali, N. Farajidavar *et al.*, "Citypulse: Large scale data analytics framework for smart cities," *IEEE Access*, vol. 4, pp. 1086–1108, 2016.

[21] M. Goldstein and S. Uchida, "A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data," *PloS one*, vol. 11, no. 4, p. e0152173, 2016.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[23] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[24] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.

[25] R. Sanchez-Iborra and M.-D. Cano, "State of the art in lp-wan solutions for industrial iot services," *Sensors*, vol. 16, no. 5, p. 708, 2016.

[26] (2017) Dash7 Alliance. [Online]. Available: http://www.dash7-alliance.org/

[27] (2017) 3GPP Low Power Wide Area Technologies, 2016 GSMA White Paper. [Online]. Available: https://www.gsma.com/iot/wp-content/uploads/2016/10/3GPP-Low-Power-Wide-Area-Technologies-GSMA-White-Paper.pdf