Received: 20 April 2023 | Revised: 21 June 2023 | Accepted: 26 July 2023

IET Wireless Sensor Systems

DOI: 10.1049/wss2.12066

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A novel system to control and forecast QoX performance in IoT-based monitoring platforms

Jose-Manuel Martinez-Caro | **Igor Tasic** | **Maria-Dolores Cano**

Department of Information and Communication Technologies, Universidad Politécnica de Cartagena, Cartagena, Spain

Correspondence

Maria-Dolores Cano.
Email: mdolores.cano@upct.es

Funding information

MCIN/AEI/10.13039/501100011033, Grant/Award Number: PID2020-116329GB-C2

Abstract

Communication architectures based on the Internet of Things (IoT) are increasingly frequent. Commonly, these solutions are used to carry out control and monitoring activities. It is easy to find cases for manufacturing, prediction maintenance, Smart Cities, etc., where sensors are deployed to capture data that is sent to the cloud through edge devices or gateways. Then that data is processed to provide useful information and perform additional actions if required. As crucial as deploying these monitoring solutions is to verify their operation. In this article, we propose a novel warning method to monitor the performance of IoT-based systems. The proposal is based on a holistic quality model called Quality of X (QoX). QoX refers to the use of a variety of metrics to measure system performance at different quality dimensions. These quality dimensions are data (Quality of Data, QoD), information (Quality of Information, QoI), users' experience (Quality of user Experience, QoE), and cost (Quality Cost, QC). In addition to showing the IoT system performance in terms of QoX in real-time, our proposal includes (i) a forecasting model for independent estimation of QoX applying Deep Learning (DL), specifically using a Long Short-Term Memory (LSTM) and time series, and (ii) the warning system. In light of our results, our proposal shows a better capacity to forecast quality drops in the IoT-based monitoring system than other solutions from the related literature.

KEYWORDS

artificial intelligence, deep learning, internet of things, performance evaluation, quality of service

1 | INTRODUCTION

Deploying new technologies, such as the Internet of Things (IoT), adds value to numerous services and applications. IoT well represents the information value loop described in ref. [1]. The sensors generate data. This data is communicated to an end device that collects data from several sources. The collected data is then analysed in search for patterns or relationships that convert data into valuable information. Finally, some actions are carried out, such as initiating or changing a physical event. The benefits of incorporating this added value into everyday applications and services [2–5] are reflected in an increase in the number of devices connected to the network, expected to be 25 billion in 2030 [6].

As crucial as deploying these IoT-based monitoring solutions is to verify their operation, and at this point, we make two observations. On the one hand, Quality of Service (QoS) metrics, that is, delay, jitter, packet losses, and bandwidth, have been widely used as objective performance evaluation metrics in communications networks [7]. These metrics are still in use, but other emerging metrics supplement them. Quality of user Experience (QoE) was first adopted as a natural evolution of QoS [7–9], incorporating new measurements to get a complete overall vision of systems' performance. Such is also the case of the quality model proposed by the authors in ref. [10], where we introduced the concept of Quality of X (QoX). QoX considers data, information, network behaviour, and user experience when evaluating or monitoring the quality of a particular system giving rise to four quality dimensions. Quality of Data (QoD) measures the quality of the raw data collected by system devices (e.g., sensors). Quality of Information (QoI) measures the quality of the information obtained after processing the data.

Quality of user Experience (QoE) enriches QoS by adding metrics relative to users and networks. Finally, Quality Cost (QC) shows the impact of the improvements in the different measured parameters according to available resources. For instance, sending more data could improve the QoI, but at the cost of using more bandwidth and consuming more energy. Therefore, QC can have a significant role in optimising resources [11]. A summary of the QoX model and its quality dimensions is shown in Table 1 and Equations (1)–(4).

$$\text{QoD} = \text{Precission} \bullet \text{Truthfulness} \bullet \text{Completeness} \tag{1}$$

$$\text{QoI} = \text{Quantity} \bullet \text{Precision} \bullet \text{Recall} \bullet \text{Accuracy} \bullet \text{Detail} \bullet \text{Timeliness} \bullet \text{Validity} \tag{2}$$

$$\text{QoE} = \text{Jitter} \bullet \text{Delay} \bullet \text{Packet Delivery Ratio} \bullet \text{Throuhgput} \bullet \text{Gateway availability} \tag{3}$$

$$\text{QC} = \text{Energy Consumption} \bullet \text{Interface Use} \tag{4}$$

On the other hand, the IoT paradigm demands new metrics for performance evaluation since the devices and systems implemented are generally different and more complex than other traditional or more studied technologies [12]. Numerous works in the related literature present IoT monitoring platforms [13–15]. However, no matter the field, the Key Performance Indicators (KPI) chosen to evaluate the performance are usually different. Some platforms focus on energy consumption or battery life, others on the accuracy of the collected data or availability, and so on. It is expected that depending on the application field, some KPIs will have more impact than others. Still, the lack of a holistic model that accommodates the diversity and influence of a wide range of KPIs or metrics in the performance evaluation hinders other tasks, such as security or standardisation. This paper proposes a system to control and forecast the performance of IoT-based platforms from a broad perspective regarding KPIs. Notably, the proposal can detect drops in the QoX of an air-pollution measurement platform and could be adapted to other monitoring systems deployed in the framework of IoT. Our proposal measures and shows QoX parameters and forecasts possible changes, generating accurate and timely warnings. Deep Learning (DL) techniques have been used, specifically, a Long Short-Term Memory with time series, to analyse independently different QoX dimensions and to predict their behaviour and performance, where the estimation is made based on the values of the last recorded measurements. The performance of our proposed system is compared with another forecasting solution from the research literature [16], showing notable results.

In summary, the contributions of this paper are the following:

TABLE 1 A summary of the Quality of X (QoX) holistic model: A case for IoT-based monitoring.

Quality dimension	Metrics	Description
Quality of data (QoD)	Precision (*P*)	Exactness of the collected measurements in every sensor
	Truthfulness (*Tr*)	Indicates the reliability degree of data resource
	Completeness (*C*)	Integrity of the sensors system
Quality of information (QoI)	Quantity (*Q*)	Information displayed and provided by LoRa devices to network server $[0 < Q < 1]$
	Precision (*P*)	Fraction of data retrieved that are relevant regard to all information obtained from sensors/networks/services $[0 < P < 1]$
	Recall (*R*)	Fraction of relevant geographic data retrieved $[0 < R < 1]$
	Accuracy (*A*)	Accuracy degree of information to the decision-maker $[0 < A<1]$
	Detail (*D*)	Complete degree of information to the processing server $[0 < D < 1]$
	Timeliness (*T*)	The information is timely for decision making $[0 < T < 1]$
	Validity (*V*)	Provided information from minimum number of IoT devices $[0 < T < 1]$
Quality of user experience (QoE)	Jitter (*J*)	Fluctuation of average delay between packets traversing the network
	Delay (*D*)	Average end-to-end delay between end devices and the server
	Packet delivery ratio (*PDR*)	Percentage of packets successfully traversing the network to the server
	Throughput (*Th*)	Average rate of *bps* received on the server during a time interval
	Gateway availability (*GW_av*)	Percentage of time that a gateway has the ethernet interface active and available to forward traffic over the traditional network to the server in a time interval period
Quality cost (QC)	Energy consumption (*EC*)	Energy required to perform an action/set of actions by all the devices under evaluation
	Interface use (*IU*)	It is binary and will take the value IU = 1 if the devices comply with the time constraint of the duty cycle in LoRa and IU = 0 if not.

(i) A novel warning method to monitor the performance of IoT-based systems is presented. The performance is assessed from a holistic perspective, including several quality dimensions, from data to cost, with the advantage of being configurable to adapt to the application field under consideration.
(ii) The operation of the proposed system is shown using collected data from a real-world scenario.
(iii) A forecasting method has been included to detect and predict malfunctioning.
(iv) The proposed complete system has been compared with a state-of-the-art solution, showing better results.

The rest of the paper is organised as follows. Section 2 presents a comprehensive review of related works. Section 3 describes the proposal and the research methodology, including the data acquisition, regression model, and warning system. The results are shown and explained in Section 4. Finally, a summary of the most important findings is included in the conclusion.

2 | RELATED WORKS

Defining a model for performance evaluation in versatile IoT-based monitoring solutions is a complex task. Numerous factors will impact, such as the number of devices, distances between the IoT devices/nodes and the communication gateway (e.g., from LoRa nodes to the LoRa gateway), assigned hardware resources, type of devices, number of software layers implemented, communication protocols etc. Any service or application constructed over IoT technology has the challenge of measuring its operation quality. For other telecommunication services like video streaming, measuring is straightforward, just by assessing network delay, packet losses, and other computer network metrics. However, there are many influencing factors for IoT-based services, so new proposals are being studied in the related literature.

In ref. [8], the authors introduced a very interesting methodology to measure quality metrics in an IoT service environment, but their focus was on the users' experience and satisfaction. A complete analysis of QoE metrics in a wireless environment and a straightforward QoE estimation were presented in ref. [17]. Similarly, in ref. [18], the authors used cognitive capabilities to optimise radio communications resources in IoT, that is, focussing on the wireless network part of the IoT-based systems. The correct selection of the most relevant KPIs was emphasised to assess a proper QoE estimation in WiFi scenarios employing supervised and unsupervised Machine Learning algorithms.

The works presented in refs. [19, 20] show how to optimise the accuracy of 5G networks minimising costs through Deep Learning techniques. In ref. [21], the authors applied a DL model based on a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN) to predict current and future values of QoE in a time series, showing the advantages of incorporating the use of DL in these scenarios. Following the same trend, a real-time model based on LSTM was introduced in ref. [22] to detect and predict anomalies in wireless networks. These anomalies were the origin of performance collapse in any device or network. As another example, Casas et al. [16] used a mobile network environment where two phases were distinguished: (i) generating a new dataset and (ii) estimating the value of quality metrics using ML techniques based on decision trees. The authors used Mean Opinion Scale (MOS) scale to evaluate system performance. Other authors [23] proposed a model to estimate QoE employing multi-line regression and using as an example an IoT scenario. Finally, the works done in refs. [24, 25] addressed the issue of cognitive abilities on IoT, following a similar approach to QoX. Although the authors in ref. [24] referred to the use of Artificial Intelligence and Machine Learning, they did not specify the application of any particular technique.

From the mentioned works, we can observe that it is critical to have a means of measuring the quality of a system operation. This is particularly important, for example, in IoT-based warning systems. When the goal is to reduce the material and economic impact of disasters and prevent loss of life, a warning system is beneficial to notify the user of a drop in quality. In other words, it will tell the user that the system's performance is decreasing; therefore, corrective actions or response plans could be executed. Nevertheless, as mentioned previously, the lack of a holistic model able to accommodate the diversity and influence of a wide range of KPIs or metrics in the performance evaluation of IoT-based monitoring hinders other tasks, such as security or standardisation.

3 | METHODOLOGY AND PROPOSAL

This section presents the research methodology followed and the description of the proposed warning system.

3.1 | Data acquisition

In our proposal, the air quality of a geographic area is monitored using IoT devices with a LoRa communication module, a Low Power Wide Area Network (LPWAN), and processing servers in the cloud. As a first step, this system is recreated using computer simulations and real data from an open air-quality dataset available in ref. [26]. Using these real data, we simulate the operation of 53 LoRa devices located across a coverage area depicted in Figure 1. Each LoRa device represents an air-quality measurement station working continuously for 360000 s (approx. 100 h). The air-quality data used in the simulation belongs to a real dataset that can be found in ref. [26] and includes the following parameters: NO, NO2, NOx, PM10, PM2.5, wind speed, temperature, wind speed, and humidity.

Each LoRa device sends the collected air-quality data to the LoRa gateway station in data packets following an exponential distribution and conforming to the limitations of the LoRa technology (duty cycle). Extensive simulations are carried out in two different scenarios. The first is an ideal scenario where the complete system behaves ideally without losses or communication delays. The second one is a scenario with failures (lossy), incorporating random losses (<10%) and network delays (<100 ms). Then, for each scenario, three different geographical environments are considered: urban, suburban, and rural, with an effect on the wireless communication channel. As a result, six test scenarios are emulated in total. For each emulated scenario, we collected the outputs in terms of quality performance QoX throughout the entire simulation, that is, the quality of raw data (QoD), the quality of the information obtained from that collected data (QoI), the quality of the communication architecture (QoE), and the quality cost (QC). Table 1 describes each quality dimension QoX and the related metrics. Equation (1)–(4) show how each quality dimension is computed. Figure 2 shows an example of the obtained quality measurements during the simulation.

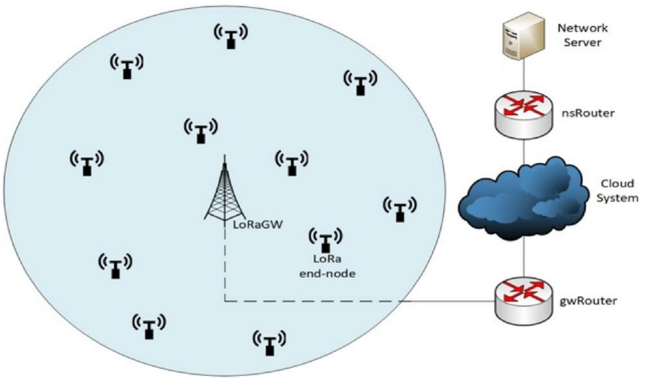


FIGURE 1 The emulated IoT-based air-quality control system. The system includes simulated devices and networks, but it uses real data. The system is composed of LoRa nodes (LoRa end-node), a LoRa gateway (LoRaGW), and a processing server (Network Server).

3.2 | Preprocessing

All outputs generated in those simulations were stored independently for each scenario and environment. This new dataset contained information about quality dimensions and KPI, namely QoD, QoI, QoE, QC, and parameters such as delay, jitter, etc.

In order to minimise the effect of the random component and maximise the performance of the proposed forecasting model and monitoring system introduced in this paper, the dataset used by our model consists of data obtained from 10 independent simulations. It is also important to note that the dataset contains a wide variety of high and low-QoX values resulting from the emulated operation using real data, so it represents well a real scenario. Later, all this information is dumped on a data processing software that allows processing a large volume of data using Python.

Next, data pre-processing eliminates failed records and groups the data. The last step in this phase is using sliding windows of size t, where t should be much smaller than the

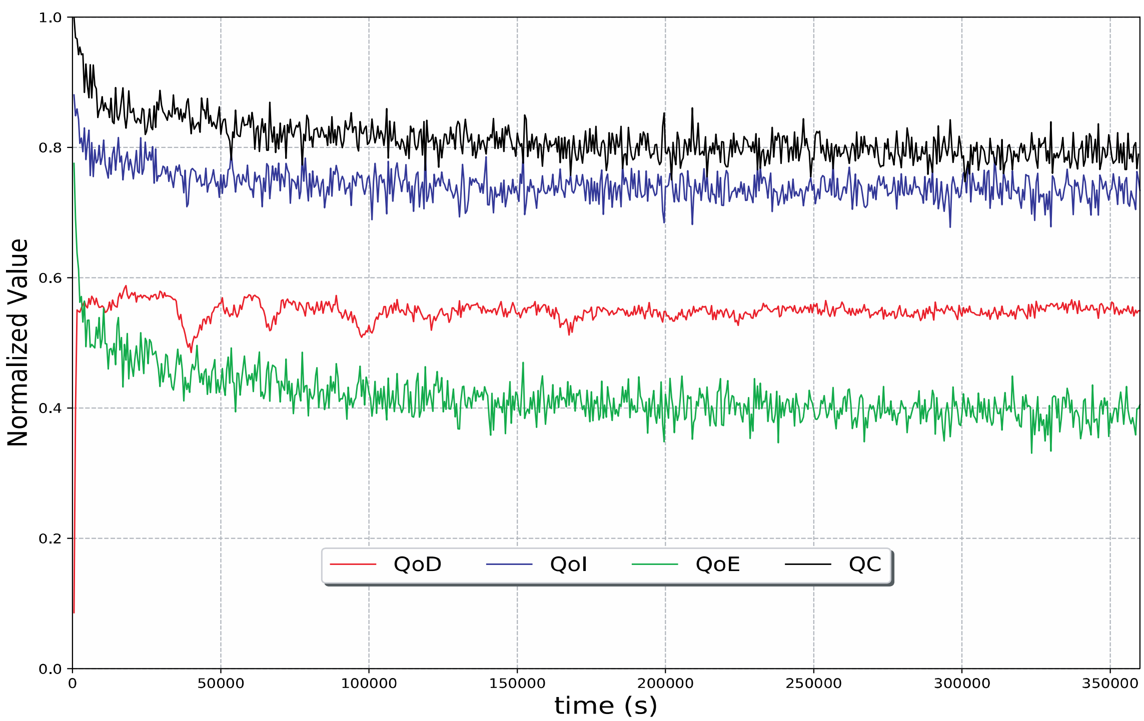


FIGURE 2 Example of Quality of X (QoX) measurements for the IoT-based air-quality system. Each line graph represents a different quality dimension (QoD, QoI, QoE, and QC) measured during the emulated operation. The closer to 1, the better the quality. Values different from 1 mean that there have been problems with any parameters related to a quality dimension as shown Table 1. For instance, it is observable that QoE is showing an average value because, in this simulation, it is impacted by delay and packet delivery ratio. More information about this quality model and its four quality dimensions can be found in [10].

total number of data collected T ($t << T$). In this case, the sliding window size is equal to 5 ($t = 5$), that is, a history of five data is required to estimate the value of the sixth one. It is important to choose a proper t because it allows (or not) the estimated value to be less (or more) conservative and adapt (or not) to sudden changes in system performance metrics. As an example (Figure 3), we need to know the values of A, B, C, D, and E in order to estimate F, and so on, until the last data is received. The smaller the window size, the more volatile the resulting prediction value is and the better it can adapt to sudden changes in our network. Once the data is organised in a time series, the data set is then divided into two parts, namely, train-set and test-set (75/25), without being shuffled.

3.3 | Proposing a Deep Learning model for forecasting

The next step is to define an independent regression model for each quality dimension QoX. One of the most popular methods for forecasting is the use of Neural Networks (NN). Neural networks are a network of interconnected nodes in which each node is responsible for simple calculations and their combination allows obtaining the desired results. Artificial NN (ANN), Convolutional NN (CNN), and Recurrent NN (RNN) are three of the most important types. RNN remember the sequence of the data and use data patterns to provide predictions, so they are used in natural language processing models and speech recognition. The main difference among RNN and other NN is the feedback loops incorporated into RNN, which facilitate data sharing among the different nodes. RNN are considered short-term memory systems, that is, the processed sequence must be relatively short for previous activations to have a relevant effect on the current prediction. However, LSTM are a particular category of RNN able to remember a relevant piece of data in the sequence and preserving it for several instants of time. Therefore, it can have long-term memory.

A	B	C	D	E		F
B	C	D	E	F		G
C	D	E	F	G		H
D	E	F	G	H		I
...	...	...	...	...		...
U	V	W	X	Y		Z

FIGURE 3 Time series with a sliding window of size $t = 5$. Real data from the first QoX measurements (A to E) is used to foresee the next QoX value (F), then real data from the next five QoX measurements (B to F) is used to foresee the next QoX value (G), and so on. This is done individually for each QoX dimension (QoD, QoI, QoE, and QC).

We propose a model based on a bi-directional LSTM with long-term memory blocks, back and forward propagation, a sliding window that adjusts the model from the training data, an optimiser based on Stochastic Gradient Descent (SGD), and a loss function based on Mean Squared Error (MSE). To manage the state and obtain the prediction of the next value, these memory blocks use gates, both input and output. Finally, forget gate is used to update the model memory. It is important to note that the model we propose is based on sequence sorting using two (bidirectional) LSTMs as input sequence as shown in Figure 4. The first one adds a copy of the sliding window data as is and the other one an inverted copy, which allows a faster and deeper learning. The training process updates the weights w of the model coefficients iteration after iteration, using SDG, optimising the weights to minimise the loss function. The proposed model has an input layer, an output layer, and a hidden layer, with the connections between the different layers being bidirectional as mentioned. The pseudocode of the proposed forecasting model is shown in Figure 5.

After the training phase, the model performance will be analysed using the test set. The R^2 score is calculated as shown in Equations (5) and (6) to assess the performance of the forecasting model. R^2 with values in the range [0,1] is also known as the determination coefficient. It determines the quality of the model to replicate the results and the proportion of variation in the results that the model can explain.

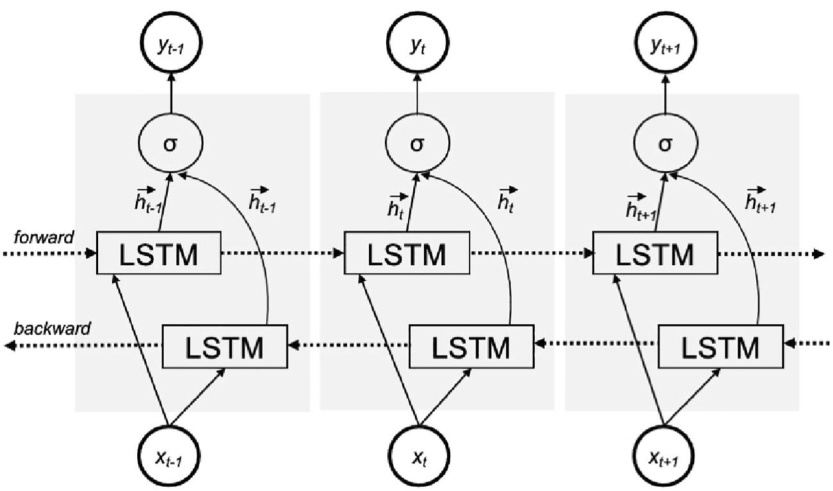


FIGURE 4 Graphical representation of the Bidirectional Long Short-Term Memory (LSTM) model (x≡input layer, h≡output sequence, y≡output layer, σ≡activation function).

Algorithm: Bi-LSTM forecasting model for QoX dimensions (QoD, QoI, QoE, and QC)

```
procedure DEFINITIONANDTRAINING (trainX[k],trainY[k], testX[k], testY[k])
    model = Sequential();
    Bidirectional(LSTM(activation='relu'));
    Dense(1);
    model.compile(optimizer='sgd', loss=mean_squared_error', metrics =
    [accuracy,F1,Recall, Precision])
    model.fit(trainX[k],trainY[k])
    return model.evaluate(testX[k],testY[k])
end procedure
procedure BI-LSTM(data,K)
    Partition the data in K subsets (DT1,…,DTK)
    for k=1 to K:
            Construct the kth training set with K-1 subsets
            Construct the kth testing set with 1 subset, different to training set.
result[k]=DEFINITIONANDTRAINING (trainX[k],trainY[k], testX[k],testY[k])
    end for
    Represent results for Loss Function, Accuracy, Precision, Recall, and F1.
```

FIGURE 5 Pseudocode of the Bi-LSTM forecasting model to predict the values of the QoX dimensions (QoD, QoI, QoE, and QC).

$$R^2 = 1 - \frac{\mathrm{MSE}(\mathrm{model})}{\mathrm{MSE}(\mathrm{baseline})} \tag{5}$$

$$\mathrm{MSE}(\mathrm{baseline}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \tag{6}$$

3.4 | Proposing a warning system for performance control

A simple yet efficient warning system is also introduced, completely configurable, that works as follows. In case the forecast predicts a d% drop in the quality performance in comparison with a previous observation time (called T_{eval} and initially set to 5 min), and if this decrease is sustained over several n consecutive T_{eval}, then a warning will be communicated to the user of the IoT-monitoring platform, together with the specific QoX dimension, metrics, and devices involved in that quality dropping. In this paper, for simplicity, it is assumed that d is equal to a 50% drop and n equals 5. Note that these values are entirely customisable and can be adapted to the specific IoT system under monitoring. Similarly, a notification warning will be generated if the estimated quality level remains identical between one T_{eval} and the next.

The proposed warning system can be seen as a classification problem that takes the value 1 when action is needed and 0 otherwise, allowing us to compare the results with other models from the related literature. A similar and interesting classification model was found in [16] for quality assessment, but that uses Mean Opinion Score (MOS) as the quality metric instead of using a more comprehensive quality model as QoX [10, 27]. The MOS scale ranges from 1-5, where 1 is the minimum value for quality (lowest quality), and 5 is the maximum (highest quality). Therefore, in contrast to our QoX model, the model presented in [16] does not allow for estimating a quality value continuously but a discrete integer value based on the KPIs collected by the simulation. So, to be able to compare that proposal with the one suggested in this paper, we focus on the notification system that they proposed, and that generated an alarm if the estimated quality value was less than 0.4 on a scale [0–1] (or 2 in the MOS scale).

3.5 | Metrics to evaluate the warning system

The robustness of the warning system will be evaluated using the confusion matrix and its four well-known components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP is the number of positives correctly classified as positive by the model. TN is the number of negatives correctly classified as negative by the model. FN is the number of positives incorrectly classified as negative, and FP is the number of negatives incorrectly classified as positive. Based on the Confusion Matrix, more advanced metrics can be calculated. Precision Equation (7) measures the quality of estimated positives concerning actual positives. Recall Equation (8) measures the number of cases classified as true positives over everything positive, and F1 Score Equation (9) is a good metric that seeks to keep estimates away from FP and FN trying to balance Precision and Recall. These advanced metrics will be employed to compare the results obtained by the warning system with our real QoX performance data and with [16].

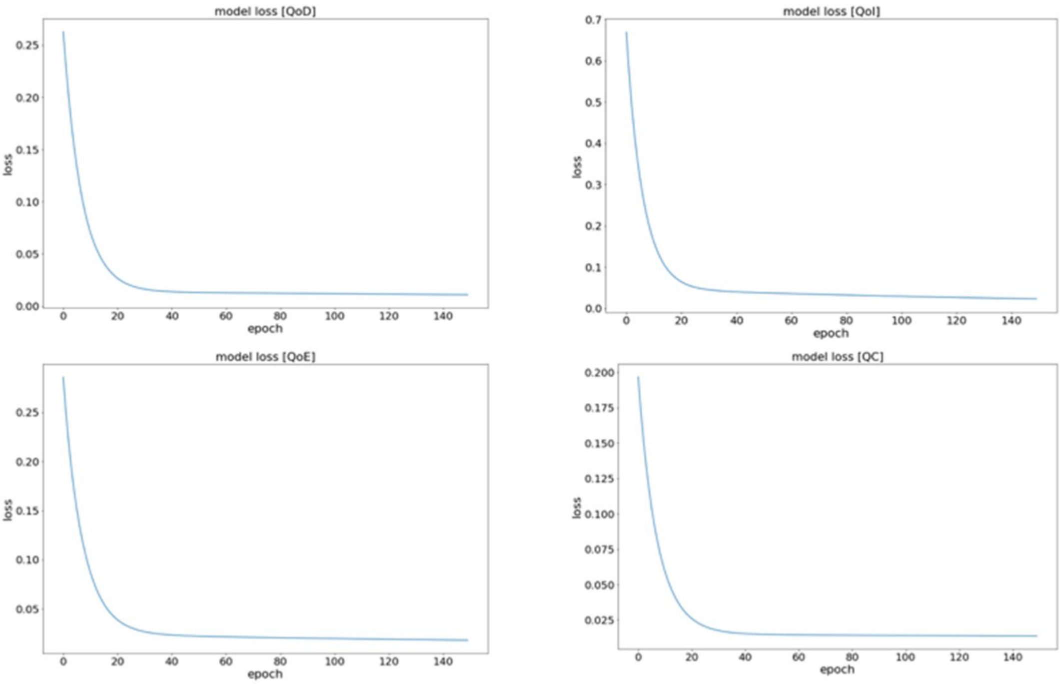
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

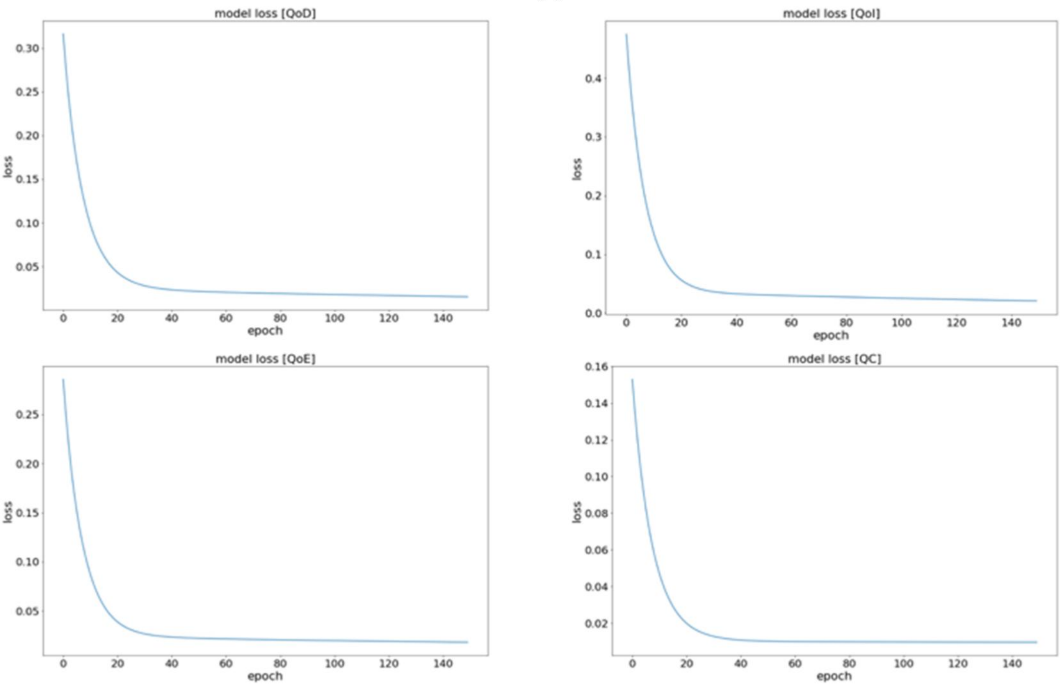
$$F1 = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \tag{9}$$

4 | RESULTS AND DISCUSSION

This section discusses the results of the proposed Deep Learning model to forecast the values of each QoX dimension (QoD, QoI, QoE, and QC). Then, the proposed warning system for performance control is compared with the contribution from ref. [16]. As mentioned before, we have carried out simulations in two different scenarios, one without losses



(a)



(b)

FIGURE 6 Loss function to evaluate the performance of the deep learning model for forecasting. (a) Loss functions in an ideal urban scenario for QoD, QoI, QoE, and QC. (b) Loss functions in a lossy urban scenario for QoD, QoI, QoE, and QC. In both cases, the model converges quickly to low values.

or communication delay and another one with random losses (<10%) and network delays (<100 ms). Then, for each scenario, three different geographical environments are considered: urban, suburban, and rural, with an effect on the wireless communication channel.

4.1 | Results of the Deep Learning model for forecasting

The loss function is obtained at the end of the training phase. We can quantify the difference between predicted and actual values by observing the loss function. In other words, it shows how good our model is in predicting the expected outcome. A high value for the loss means our model performed poorly. A low value for the loss means our model performed very well. We can see in Figure 6 that the loss functions are quite similar. Figure 6a represents the loss function in the ideal scenario described above for QoD, QoI, QoE, and QC. Similarly, Figure 6b represents the loss function in the lossy scenario for QoD, QoI, QoE, and QC. It can be deduced from Figure 6 that the model converges reasonable quick and well.

Once training is finished, we see the results of the test phase. Again, using time series, our model estimates the QoX value and compares it to the real one, but this time, using the test set. Throughout the test set, there are significant performance drops of variable duration over the recorded data, and our proposed model allows us to estimate performance drops correctly, including short-duration cases. This is due to the sliding window size used in the time series ($t = 5$). Because it is essential to consider multiple metrics when evaluating the performance of a deep learning regression model, we also use the R^2 score to measure and compare the performance of the deep learning model for each QoX dimension in each emulated scenario of the IoT-based air-quality monitoring system. The closer to 1 R^2 is the better the result. As shown in Figure 7, the best forecasting performance is achieved for the QoI dimension in both the ideal and lossy environments, with an approximate R^2 value of 0.7. Figures 8 and 9 include two examples with the forecast for QoI and QoE and the real values respectively. A reasonable forecast is achieved for the other quality dimensions, QoD and QC. Therefore, on average, the performance obtained by the proposed forecasting model is good.

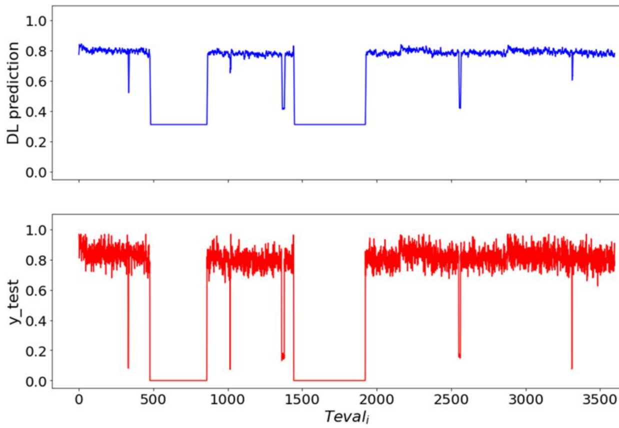


FIGURE 8 Example of the results obtained when forecasting QoI (above) versus the real values (below), with $R^2 = 0.75$. It can be seen that the drops in quality (QoI dimension) are well predicted by the model, both the large and the small drops. ($T_{\text{eval}} = 5$ min).

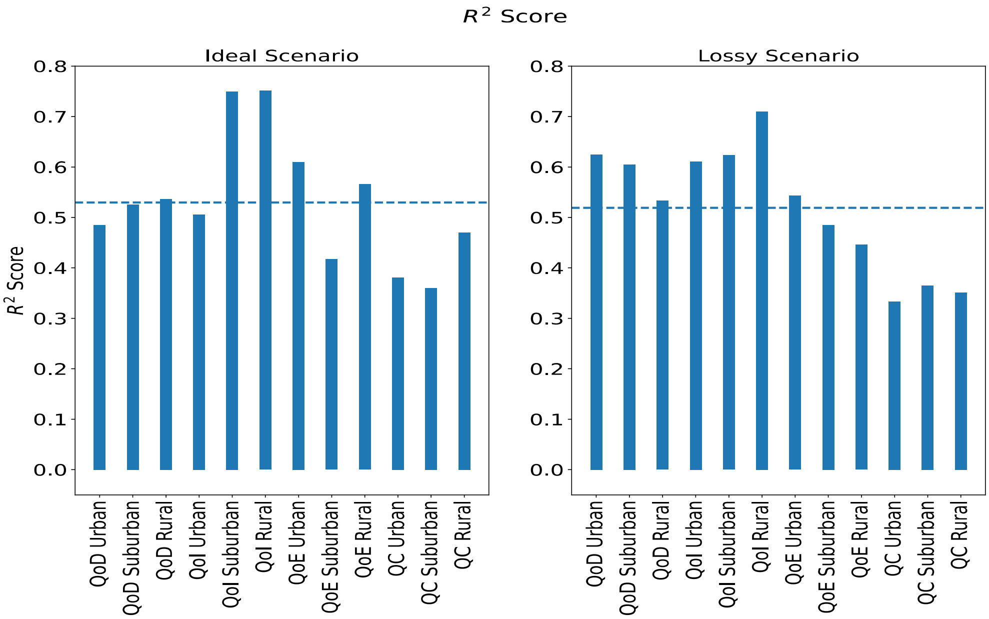


FIGURE 7 Results in terms of R^2 of our deep learning QoX forecasting model under different emulation conditions. The graph on the left corresponds to the ideal scenario. The graph on the right corresponds to the scenario with delays and packet losses. Then, for each scenario, the four quality dimensions (QoD, QoI, QoE, and QC) are displayed in three different environments (urban, suburban, and rural).

4.2 | Results of the warning system for performance control

In our warning system, there will be a notification when two conditions are met. First, if there is a sharp drop in the performance estimation of QoX dimensions, and second, if this drop is maintained for a specific time interval. Particularly, a warning will be communicated to the user of the IoT-monitoring platform, together with the specific QoX dimension, metrics, and devices involved in that quality dropping in case the forecast predicts a 50% drop in the quality performance in comparison with the previous observation time (T_{eval}), and if this decrease is sustained over five consecutive T_{eval}. As mentioned, these values are entirely customisable and can be adapted to the specific IoT system under monitoring.

Figures 10 and 11 illustrate the operation of the warning system. In these graphs, the y-axis is a boolean value, 0 or 1. 1 means that an action is required due to a drop in the quality dimension under control, whereas 0 means that no action is required because the quality level is kept.

As we have seen in the related literature, some articles address issues similar to this case study, but to the authors' knowledge, none apply a holistic and multi-dimensional model. In order to make a comparison with other works, we will focus on the case of QoE monitoring, evaluation, and estimation. In particular, our warning system will be compared with the one introduced in ref. [16]. In that paper, the authors stated that obtaining data from all levels of the protocol stack was necessary to have an accurate estimation. Regarding the estimation, they compared different ML models to predict the QoE value and obtained the best performance with the C4.5 classification model based on decision trees. This model estimates the MOS value based only on the following parameters: gateway availability, jitter, delay, throughput, quantity, packet delivery rate, and recall, which were also part of our QoX model (see Table 1).

A priori, the two forecasting models would not be comparable since the problem is addressed by employing two different solutions, one with a classification algorithm [16] and the other with a regression algorithm (our proposal). Therefore, to compare our proposal with ref. [16], we replicate the model presented in ref. [16] to obtain the forecast QoE values. Then, we apply the same warning system but either using the forecast QoE data obtained with the model from ref. [16] or our forecast QoE data obtained using our forecasting model. The results of the notification system with either input are compared to the real values using the confusion matrix and the Recall, Precision, and F1 metrics. This comparison is only performed for the QoE dimension since it is the only one used by the authors in ref. [16]. Each subplot in Figure 12 represents a scenario (ideal or lossy) and an environment (urban, suburban, and rural). The graphs in the upper row show the associated performance for each

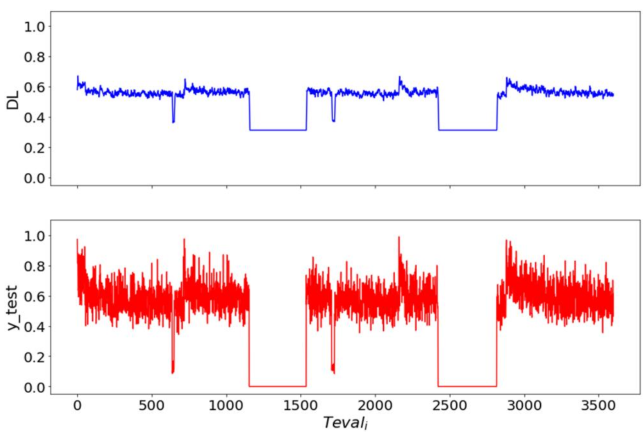


FIGURE 9 Example of the results obtained when forecasting QoE (above) versus the real values (below), with $R^2 = 0.566$. It can be seen that the drops in quality (QoE dimension) are well predicted by the model, both the large and the small drops. ($T_{\text{eval}} = 5$ min).

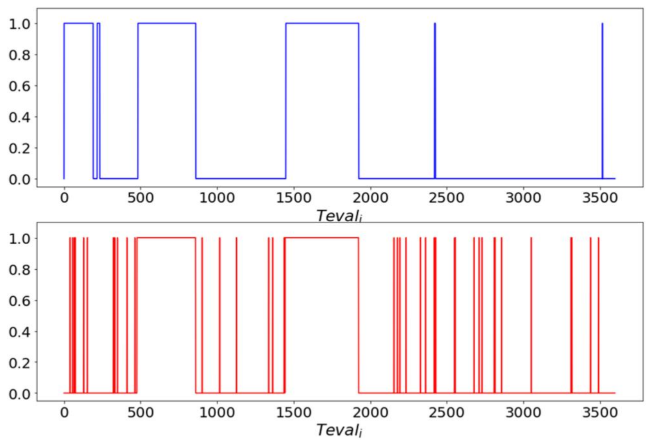


FIGURE 10 Example of the results obtained by the warning system for the quality dimension QoI. The lower graph represents the real QoI performance; 1 means that during five consecutive T_{eval} the value of QoI has decreased 50%. The upper graph shows the output of the warning system based on predictions, with Recall = 0.95, Precision = 0.8, and $F1 = 0.87$. ($T_{\text{eval}} = 5$ min).

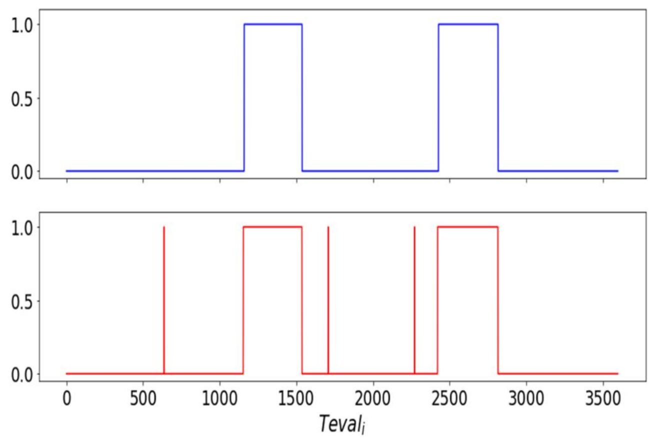


FIGURE 11 Example of the results obtained by the warning system for the quality dimension QoE. The lower graph represents the real QoE performance; 1 means that during five consecutive T_{eval} the value of QoE has decreased 50%. The upper graph shows the output of the warning system based on predictions, with Recall = 0.98, Precision = 1, and $F = 0.99$. ($T_{\text{eval}} = 5$ min).

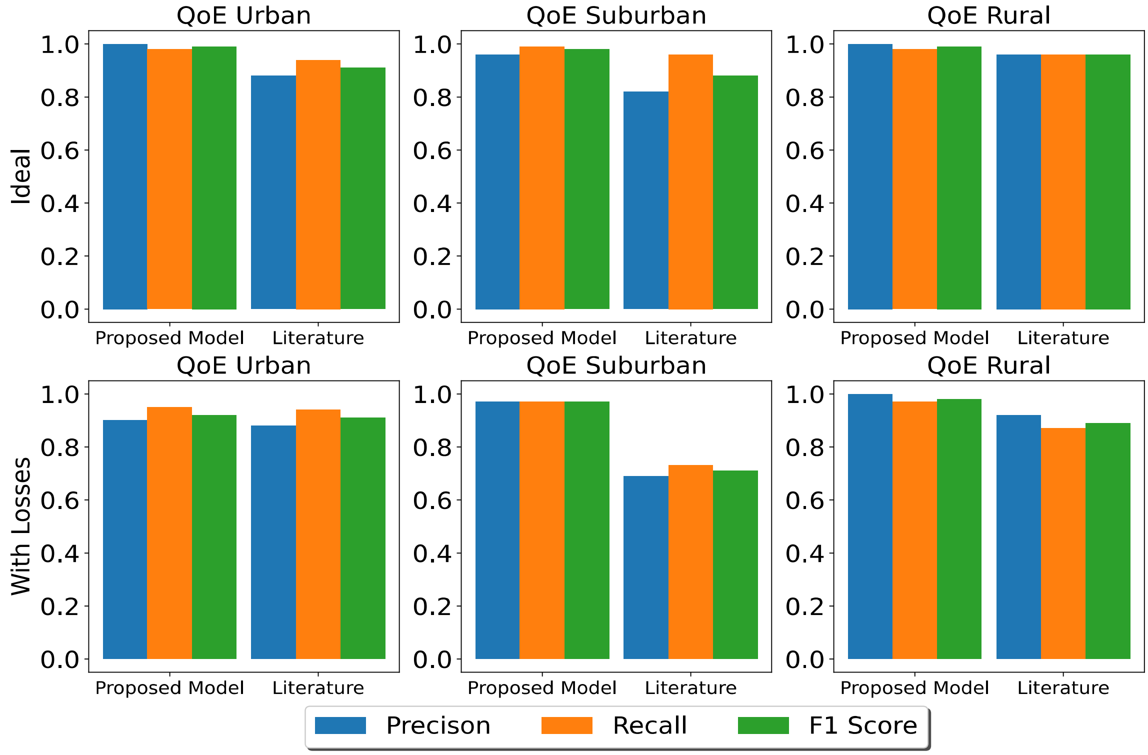


FIGURE 12 Comparing the performance of our warning system with the performance of ref. [16]. Results are shown in terms of Precision, Recall, and F1-Score for the two scenarios (ideal and lossy) and the three environments (urban, suburban, and rural). The comparison has been made using the QoE quality dimension.

environment under an ideal scenario, while the lower graphs show it for a lossy scenario. The urban, suburban, and rural environments are represented from left to right. In all scenarios and environments, our warning system performs better in terms of Precision, Recall, and F1-score.

Both proposals' results are similar in some cases, such as in rural and urban environments. The differences are more significant in a suburban environment, especially in a lossy scenario that is closer to reality. For instance, Figure 13 compares model estimation for QoE in the lossy suburban environment, and Figure 14 illustrates warning system operation comparing the method proposed in ref. [16], our proposal, and the real values. Considering all the evaluated scenarios, the average improvement obtained with our proposal is 11.4% in Precision, 7.3% in Recall, and 14% in F1 Score.

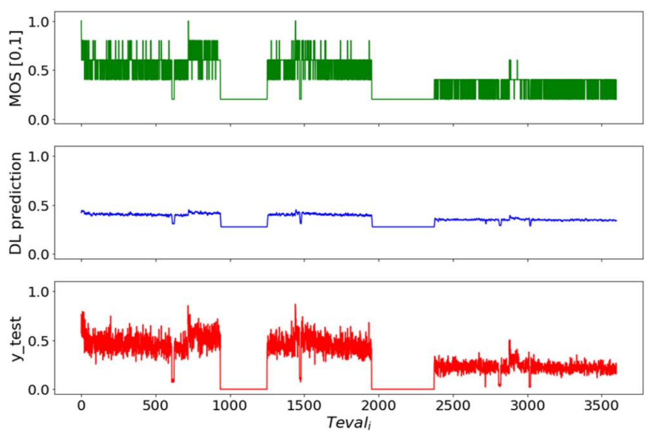


FIGURE 13 Comparison of forecasting model with respect to training values for QoE in lossy suburban environment. At the top, it is represented the MOS obtained using the model from ref. [16], in the middle, our forecasting model, and below the real QoE measurements.

5 | CONCLUSION

The development and deployment of IoT technology are advancing rapidly. New applications, devices, and systems are being introduced because of their higher efficiency. While introducing improvements, it is essential to be able to measure the level of quality of these services and applications. An unexpected drop in the performance of these solutions, which are often used to monitor risk situations or environments, could have dramatic consequences. We have seen that several models have been proposed to measure the quality of IoT-based systems. However, just a few introduce a holistic approach that can encompass all the features inherent to this type of IoT monitoring services, and none present practical results. For the first time, to the authors' knowledge, this paper addresses the quality of an IoT system as the conjunction of several quality dimensions, called QoD, QoI, QoE, and QC, and in general,

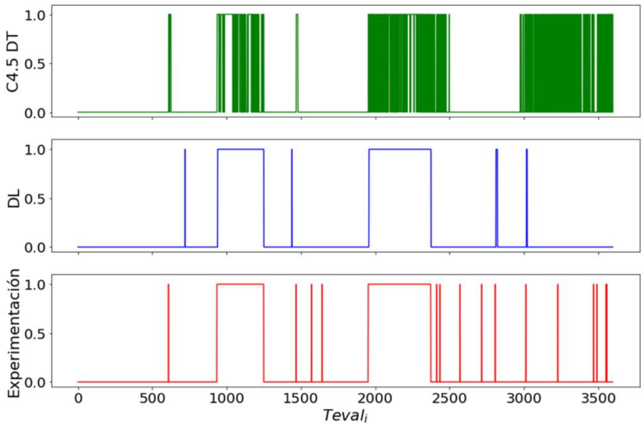


FIGURE 14 Comparison of the warning system to detect quality drops for QoE in a lossy suburban environment. At the top, it is represented the model from ref. [16] (Recall = 0.73, Precision = 0.69, F1 0.71; TP = 549, TN = 2601, FP = 202; False Negative (FN) = 247), in the middle, the output of our warning system (Recall = 0.97, Precision = 0.97, *F1* = 0.97; TP = 726, TN = 2828, FP = 25; FN = 20), and below the QoE warnings based on real values.

QoX. An IoT-based air quality monitoring service has been emulated using real data, including a deep learning model using LSTM to forecast the behaviour of the four quality components and a warning system to communicate sustained quality drops. The system has been compared with another method from the related literature, showing better performance and a broader spectrum of use. We plan to improve the models in future work and conduct new experimental tests in real environments.

AUTHOR CONTRIBUTIONS

Jose-Manuel Martinez-Caro: Conceptualisation; data curation; formal analysis; investigation; methodology; software; validation; visualisation; writing – original draft. **Igor Tasic**: Conceptualisation; validation; visualisation; writing – review & editing. **Maria-Dolores Cano**: Conceptualisation; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualisation; writing – original draft; writing – review & editing.

ACKNOWLEDGEMENTS

This work was supported by Grant PID2020-116329GB-C22 funded by MCIN/AEI/10.13039/501100011033.

CONFLICT OF INTEREST STATEMENT

There are no conflict of interests.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the references.

ORCID

Maria-Dolores Cano https://orcid.org/0000-0003-4952-0325

REFERENCES

1. Holdowsky, J., Mahto, M., Raynor, M.: Inside the Internet of Things (IoT) A Primer on the Technologies Building the IoT. https://www2.deloitte.com/us/en/insights/focus/internet-of-things/iot-primer-iot-technologies-applications.html. Accessed March 2022
2. Tsolkas, D., et al.: A survey on parametric QoE estimation for popular services. J. Netw. Comput. Appl. 77, 1–17 (2017). https://doi.org/10.1016/j.jnca.2016.10.016
3. Al-Turjman, F. (ed.) Performability in Internet of Things. Springer, Cham (2019)
4. Guillen-Perez, A., Cano, M.D.: Intelligent IoT systems for traffic management: a practical application. IET Intell. Transp. Syst. 15(2), 273–285 (2021). https://doi.org/10.1049/itr2.12021
5. Tasic, I., Cano, M.D.: Sparking innovation in a crisis: an IoT sensor location-based early warning system for pandemic control. Appl. Sci. 12(9), 1–15 (2022). https://doi.org/10.3390/app12094407
6. Statista The Statistics Portal: Number of Internet of Things (IoT) Connected Devices Worldwide from 2019 to 2030. https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide. Accessed March 2022
7. Tu, W.: Data-driven QoS and QoE management in smart Cities: a tutorial study. IEEE Commun. Mag. 56(12), 126–133 (2018). https://doi.org/10.1109/mcom.2018.1700870
8. Suryanegara, M., et al.: A 5-step framework for measuring the quality of experience (QoE) of Internet of Things (IoT) services. IEEE Access 7, 175779–175792 (2019). https://doi.org/10.1109/access.2019.2957341
9. Park, J., Chung, K.: Resource prediction-based edge collaboration scheme for improving QoE. Sensors 21(24), 8500 (2021). https://doi.org/10.3390/s21248500
10. Martinez-Caro, J.M., Cano, M.D.: A holistic approach to evaluate the performance of applications and services in the Internet of Things. Int. J. Commun. Syst. 34(2), 1–19 (2021)
11. Liu, C.H., et al.: Toward QoI and energy-efficiency in internet-of-things sensory environments. IEEE Trans. Emerg. Top. Comput. 2(4), 473–487 (2014). https://doi.org/10.1109/tetc.2014.2364915
12. Mutichiro, B., Tran, M.N., Young-Han, K.: QoS-based service-time scheduling in the IoT-edge cloud. Sensors 21(17), 5797 (2021). https://doi.org/10.3390/s21175797
13. Andrianto, H., et al.: Performance evaluation of IoT-based service system for monitoring nutritional deficiencies in plants. Inf. Process. Agric, 1–19 (2021)
14. Mörth, O., et al.: IoT-based monitoring of environmental conditions to improve the production performance. Procedia Manuf. 45, 283–288 (2020). https://doi.org/10.1016/j.promfg.2020.04.018
15. D'Angelis, O., et al.: IoT architecture for continuous long term monitoring: Parkinson's Disease case study. Internet of Things 20, 1–19 (2022)
16. Casas, P., et al.: Predicting QoE in cellular networks using machine learning and in-smartphone measurements. In: 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6 (2017)
17. Aloqaily, M., et al.: A multi-stage resource-constrained spectrum access mechanism for cognitive radio IoT networks: time-spectrum block utilization. Future Generat. Comput. Syst. 110, 254–266 (2020). https://doi.org/10.1016/j.future.2020.04.022
18. Ibarrola, E., et al.: QoE enhancement in next generation wireless ecosystems: a machine learning approach. IEEE Commun. Stand. Mag. 3, 63–70. https://doi.org/10.1109/mcomstd.001.1900001
19. Schwarzmann, S., et al.: Accuracy vs. Cost trade-off for machine learning based QoE estimation in 5G networks. In Proc. IEEE Int. Conf. Commun, 1–6 (2020)
20. Shubyn, B., et al.: Deep learning based adaptive handover optimization for ultra-dense 5G mobile networks. In: In Proc. 15th Int. Conf. Adv. Trends Radioelectron. Telecommun. Comput. Eng. TCSET, pp. 869–872 (2020)

21. Lopez-Martin, M., et al.: Deep learning model for multimedia quality of experience prediction based on network flow packets. IEEE Commun. Mag. 56(9), 110–117 (2018). https://doi.org/10.1109/mcom.2018.1701156
22. Sui, T., et al.: A real-time hidden anomaly detection of correlated data in wireless networks. IEEE Access 8, 60990–60999 (2020). https://doi.org/10.1109/access.2020.2984276
23. Li, L., Rong, M., Zhang, G.: An Internet of Things QoE evaluation method based on multiple linear regression analysis. In: Proc. 10th International Conference on Computer Science & Education (ICCSE), pp. 925–928 (2015)
24. Minovski, D., Ahlund, C., Mitra, K.: Quality of IoT experience in autonomous vehicles. IEEE Internet Things J. 7(5), 3833–3849 (2020). https://doi.org/10.1109/jiot.2020.2975418
25. Deebak, B.D., et al.: IoT-BSFCAN: a smart context-aware system in IoT-Cloud using mobile-fogging. Future Generat. Comput. Syst. 109, 368–381 (2020). https://doi.org/10.1016/j.future.2020.03.050
26. Euskadi Government: Euskadi Air Quality (2018). http://opendata.euskadi.eus/catalogo/-/calidad-aire-en-euskadi-2018. accessed April 2022
27. Martinez-Caro, J.M., Cano, M.D.: IoT system integrating unmanned aerial vehicles and lora technology: a performance evaluation study. Wireless Commun. Mobile Comput. 2019, 1–12 (2019). https://doi.org/10.1155/2019/4307925

How to cite this article: Martinez-Caro, J.-M., Tasic, I., Cano, M.-D.: A novel system to control and forecast QoX performance in IoT-based monitoring platforms. IET Wirel. Sens. Syst. 13(5), 178–189 (2023). https://doi.org/10.1049/wss2.12066