BUILDING PREDICTIVE MAINTENANCE FRAMEWORK FOR SMART ENVIRONMENT APPLICATION SYSTEMS

Attila Katona^a & Peter Panfilov^b

^aUniversity of Applied Sciences Technikum Wien, Höchstädtplatz 6, 1200 Wien, Austria

^bNational Research University – Higher School of Economics, Myasnitskaya St. 20, Moscow 101000, Russian Federation

This Publication has to be referred as: Katona, A[ttila] & Panfilov, P[eter] (2018). Building Predictive Maintenance Framework for Smart Environment Application Systems, Proceedings of the 29th DAAAM International Symposium, pp.0460-0470, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-9679, Vienna, Austria

DOI: 10.2507/29th.daaam.proceedings.068

Abstract

Predictive maintenance is a powerful maintenance strategy that makes it possible to significantly reduce operation and maintenance costs of public, commercial and industrial environments. It is a complex data-driven process, which tries to forecast future states of company assets. On one hand it prerequisites condition monitoring of components on machine level. On the other hand it demands the integration of the collected data with other management information systems. Digitization and especially the advent of big data science bring along promising opportunities to create effective smart monitoring and predictive maintenance applications. The aim of this research is to examine the possibilities of a predictive maintenance framework based on the design principles of Industry 4.0 and recent developments in distributed computing, Big Data and Machine Learning. It introduces numerous enabling technologies such as industrial Internet of things, standardized communication protocols, as well as edge and cloud computing. Moreover, it takes a deeper look at data analytical techniques and tools, and analyses performance of well-known machine learning algorithms. Paper proposes architecture of a predictive maintenance framework based on existing software and hardware solutions. As a proof of concept, a real-life smart heating, ventilation, and air conditioning (HVAC) application system is created and tested to demonstrate the possibilities of the proposed PdM framework.

Keywords: predictive maintenance; distributed computing; big data analytics; smart environment

1. Introduction

Many industrial, commercial and public applications, such as manufacturing systems, financial institutions, power stations or medical systems, have to work 24/7. Some of these systems may operate in distant or harsh environments, such as wind farms, oil rigs, etc. Every downtime of these systems can result in enormous monetary losses and in case of errors high repair costs. In order to improve availability and reliability of such systems an appropriate operation and management (O&M) program is to be introduced. If companies want to stay on the competitive edge and improve market position, they need to reduce their O&M costs through integration of last developments in smart environments including smart metering/monitoring, smart maintenance, and smart governance solutions into their O&M programs.

There are some well-established maintenance strategies with their advantages and disadvantages. On one hand, reactive strategies represent lower O&M costs. However, in case of a failure, the whole system has to be shut down and the erroneous component needs to be repaired or replaced. On the other hand, preventive strategies mean higher O&M costs, but they can in many cases prevent component breakdowns. Predictive maintenance (PdM) represents a promising, state-of-the-art and intermediate solution based on failure detection and component/system status prediction. Proactive maintenance adds to the PdM capacities to reconfigure and improve its own efficiency on the basis of root cause analysis techniques. These help to identify the source of problem and to find a most efficient way of maintenance.

The adoption of PdM represents an enormous business opportunity for enterprises that wish to take advantage of cutting-edge technologies of digitization [1]. According to a company ranking made by IoT Analytics, the top five companies enabling PdM are IBM, SAP, Siemens, Microsoft and GE [2]. Just to mention some real-life examples, Siemens has successfully rolled out a PdM pilot for Deutsche Bahn in 2016 and GE has used its Asset Performance Management system to improve BP's oil and gas production. Dr. Stefan Ferber claims that the goal of PdM is to prognosticate machine errors before they arise and subsequently reduce repair costs [3].

In order to implement an effective component/system "health" monitoring system, it is necessary to introduce a PdM framework. There are numerous requirements for such a system; one of the most important is the communication and integration requirements of both data from machine (shop floor) level and from upper enterprise (management) level. This can be achieved by the utilization of well-known industrial protocols, such as OPC UA, MQTT, etc. Moreover, in order to achieve sustainable business goals, it is unavoidable to take advantage of a whole stack of new technological instruments including Smart meters/sensors and Industrial Internet of Things (IIoT), edge computing (EC) and cloud computing (CC), Big Data analytics or machine learning (ML) [4].

The goal of this research is to propose feasible architecture and an implementation roadmap of a PdM framework, which could be applied in different industrial environments and for complex cyber-physical systems of various scales, from smart buildings, through smart enterprises to even smart cities. During the research and the development, the authors have investigated how technological innovations such as IIoT, the combination of EC and CC the ML algorithms and Big Data applications can support and improve decision making processes in modern smart maintenance practice. The authors believe that the EC is relevant topic in the era of IIoT, and a special focus is made on this approach. The research strategy is based on thorough study of literature, analysis of reference projects as well as the application of the above mentioned cutting-edge technologies. Subsequently, a real-life smart "heating, ventilation, and air conditioning" (HVAC) application system is created to demonstrate possibilities of the proposed PdM framework.

The rest of the paper is organized as follows. To begin with, Section 2 introduces maintenance management problem area and existing solutions based on the Computerized Maintenance Management System (CMMS) concept. A simple control-loop model of the PdM process and reference architecture of a "typical" PdM framework is discussed. In Section 3, new architecture of a PdM framework is proposed as an extension to the reference architecture in Section 2 that comprises edge and cloud computing technologies in data-driven distributed computing environment and facilitating big data approaches to predictive data analytics. A PdM process model and its implementation roadmap accompany the architecture details with description of all needed steps to complete the development process. In Section 4, an experimental setup and results are discussed. It presents a prototype PdM application system for the "Smart building" demonstrator using simple HVAC application as a test case, which integrates state-of-the-art EC and CC technologies and anomaly detection executed at the edge and a data reporting and visualization facilities in the cloud. Finally, section 5 summarizes the whole research project and suggests some future developments.

2. Smart maintenance management

Maintenance management in modern industrial environment is a crucial and expensive task. The goal of it is to prolong equipment lifetime, prevent unplanned downtimes of any facilities and increase overall safety. Badly calibrated component could go wrong in a much more shorted time. Unreliable operation of manufacturing systems represents a risk to lose existing and potential clients. An incorrectly functioning power plant could explode and claim human lives. Therefore, the application of an appropriate maintenance management system is essential in every modern industrial organization. Maintenance management is part of the facility management. It aims at establishing operational processes that improve safety and performance of companies and increase return on investment (ROI). Furthermore, a well implemented cost-effective maintenance strategy can contribute to the sustainable development of the company.

With the help of well implemented and sustainable maintenance strategies, enterprises can preserve their market competitiveness, decrease O&M costs and increase profit. According to the definition of the European Federation of National Maintenance Societies, maintenance comprises combination of all technical and the associated administrative, management and supervision actions aiming at preserving or restoring specific condition of a piece of equipment, machine, or system in order to extend its lifetime [5]. Maintenance management is a collection of complex processes. It has a direct influence on internal processes such as production, quality assurance, and environment improvement but also on risk analysis and safety of people and organizations. Considering the versatility and complexity of such a smart environments as modern manufacturing plants it is reasonable to introduce smart maintenance strategies [6].

Predictive maintenance is an important enabler of Industry 4.0 concept. It is based on Computerized Maintenance Management System (CMMS) concept that takes advantage of state-of-the-art technological innovations [7]. CMMS coordinates all activities related to the availability, productivity and maintainability of cyber-physical systems (CPSs).

CMMS functioning is a data-driven process on the basis of condition monitoring. The measurements of system's component states in a complex CPS are conducted on regular basis or in real-time. High degree of automation is another crucial characteristic of such systems. All steps of a computerized maintenance process happen without or with minimal human involvement. Such approach prevents or minimises effects of human mistakes. To implement it, complex CMMS are required, which work fully autonomously, and are able to collect, store and analyze data continuously.

In order to predict future failures, downtimes or other critical events, it is necessary to analyze previously collected historical data. At the same time, it is required to monitor the current status of the assets by collecting real-time data. Application of mathematical and statistical methods enables the smart maintenance system to detect where, when and why a component is likely to fail.

In case of predictive maintenance, the equipment gets repaired or replaced well before any failure occurs [8]. PdM results in higher availability and reliability of the equipment. It also improves product quality and helps to optimize asset management. Some of the most important features of PdM are the Time to Failure (TTF) prediction and the Remaining Useful Life (RUL) prediction. The TTF prediction is the amount of time, which a component is expected to last in operation. The RUL of a component is the estimated timespan, until it remains capable to serve its intended purposes. The estimation of TTF and RUL is really a challenging task. However, it's been proven that these parameters are applicable to describe rotating machineries, such as engines, pumps and fans.

Modern distributed computing architectures and data-driven application systems offer dramatic improvements in efficiency of maintenance service. An important challenge of a PdM is to bring together technologies from different application domains including Big Data, Internet of Things, machine learning and deep learning, real-time simulation and artificial intelligence, advanced user interfaces of virtual and augmented reality (AR/VR) [9]. The key factor of the PdM system success is an interface to other enterprise systems such as ERP.

The operation of a simplified PdM can be represented with the four-state control loop [3]. As a first step, raw data is collected from the physical world by sensors. On a second step, the cleansed data is stored in a central data base. Then, the data is analyzed and interpreted. Finally, the next maintenance service activity is planned.

The reference architecture of the PdM framework has been proposed in [10] as a service-oriented architecture, which comprises various functionalities executed on different layers. The PdM system itself is located between lower-level manufacturing domain and upper-level enterprise systems domain. On one hand, at the manufacturing level, it offers interfaces to vendor specific systems, such as machines, PLC, etc. On the other hand, it retrieves complementary (historic) data from enterprise information systems for predictive analytics.

The interfaces for data acquisition on manufacturing domain have a hierarchical structure. This approach is beneficial, since the supported third-party systems have different maturity levels. For instance, some of the connected components provide only unprocessed raw data whereas other more complex ones already have predictive capabilities too. The heterogeneity of connected systems requires the support of various standards and protocols.

In addition, the reference framework suggests the application of a generalized prediction process, which is based on the execution of modelling steps. Depending on the maturity level of the manufacturing domain components, different modelling steps get executed. The framework comprises system, dynamic and predictive modelling techniques. The centralized model framework represents an abstract, integrated and unified view of them.

In order to combine the data stemming from manufacturing domain and ERP system, there is a need for a consolidation and fusion functionality. At this layer, several predictions related to the observed component are combined to achieve higher accuracy in future forecasting. This functionality is supported by an analytical warehouse, which integrates all the relevant data, such as ERP master data and manufacturing domain's model and process data. This layer is responsible for notifying enterprise systems about potential failures of components before their occurrence.

3. Predictive maintenance framework

The realization of a PdM framework based on big data and distributed data-driven computing is not a trivial task. It requires a lot of expertise and interdisciplinary knowledge. Moreover, it demands from enterprises organizational as well as technical prerequisites. The PdM development team needs to fully understand the problems of the target business or industrial smart environment. Moreover, they also need to master the full technological stack required to realize a secure and cloud based IIoT architecture. The goal of this research is to support industry players, who are considering the introduction of such a PdM framework to their industrial and business needs. In the following subsections, our own architecture of a PdM framework is described as an extension to the reference architecture that comprises edge and cloud computing technologies and facilitating big data approaches. A possible implementation roadmap will accompany the architecture details with description of all needed steps to complete the development process. Finally, the functionalities of a ready to use PdM framework are demonstrated.

3.1. Architecture of PdM framework

Our approach to architecting a PdM framework, on one hand, utilizes some of the ideas of the reference architecture in [11], particularly, in that the PdM is positioned exactly in between manufacturing domain and enterprise information system domain.

From the computing and networking point of view, it is a distributed data-driven application system that follows a client-server model. It offers prediction models which are based on well-known big data analytics solutions. Moreover, it is capable of consolidating and integrating data arriving from low-level equipment layer and high-level enterprise information systems layer. On the other hand, it facilitates state-of-the-art technological (software and hardware) stack with embedded big data processing capabilities, including those SW/HW parts on the server side (e.g., Pharos Navigator or IBM Watson™ IoT Platform) and on the client side (e.g., Apache Edgent, smart sensors and actuators, CPSs, enterprise information systems (ERP, CRM), data visualization and reporting systems, etc.). Figure 1 shows the principal architecture of the proposed PdM framework.

Fig. 1. High-level architecture of a PdM framework

On the server side, it is suggested, that the Pharos Navigator from Golem IMS GmbH can be used as the main building element of the server-side system. It has an event-based behavior and functions as Software as a Service (SaaS). Its tasks involve but are not limited to data archiving, data analytics and support of external interfaces. In most cases it is beneficial to launch it in a cloud environment, since the cloud infrastructure supports the concept of on-demand computing. The computational and storage resources are made available, when they are needed. This approach is suitable for handling big data streams from numerous data sources. The software running on server is responsible for the orchestration and archiving of the incoming data streams. Moreover, it executes deep learning algorithms and in case of any suspicious behaviors, it alerts the connected systems. It supports various industrial protocols and standards and it offers interfaces for several interfaces. This approach results in higher availability, scalability and security. For standardized visualization and configuration purposes, Pharos Navigator offers a web-based user interface. It has a hierarchical structure and allows getting a deeper insight into the ongoing processes of the PdM activities.

On the client side, the proposed architecture supports almost any number of connected clients. The clients can be globally distributed. The communication between these clients to the centralized Pharos Navigator server happens via Internet, with the help of well-known industrial protocols, such as OPC UA, MQTT, IoTivity, etc. The clients are divided into two types: providers and consumers.

Providers are the sources of data used for further predictive analytics. As already mentioned, data can be collected from systems with various maturity levels. It can be raw data directly from smart sensors, actuators or CPSs. Then, it can be already cleansed and processed data from SCADA systems. Also it can be data provided by edge systems, such as Apache Edgent. In this case the data is processed in real-time and it can contain prediction information too. The data collected from different subsystems, such as smart sensors, CPS, actuators, SCADA are sent to an IoT gateway of Apache Edgent. All these components are within a local network, in a “demilitarized” zone. Additionally, the data can originate from high-level enterprise information systems, such as ERP. It is the complementary (historical) data for predictive analytics purposes.

Consumers are the clients of Pharos Navigator application that receive data or information from the server side. On one hand, enterprise systems act as consumers too, since they integrate the results of predictive analytics into their internal data structures. This allows triggering enterprise-wide events, change business processes such as purchase, storage management, maintenance management. On the other hand, consumers can be data visualization and reporting application systems. They range from simple reporting and alarm systems, through mobile devices and applications, to smart devices with VR/AR capabilities.

3.2. PdM process

The proposed PdM system aims at a fully automated approach to monitor company assets analyze their status and trigger subsequent maintenance actions in case of a suspicious or anomalous behavior. The whole system is data-driven and represents a much more effective, customized application to prolong the lifetime of company's properties and assets, increase product quality, and optimize maintenance operations. Ultimately, it brings better customer satisfaction. Figure 2 shows the whole chain of the proposed PdM process.

Fig. 2. PdM process model

As a first step process data is collected from many different sources. These are smart sensors, actuators, CPSs, SCADA systems, and other.

The collected data are transmitted to a local Apache Edgent platform. At the edge, the erroneous and incomplete data get cleansed. Also, some important filtering functions are executed on the incoming data streams. Subsequently, the preliminary edge data analysis is executed with real-time constraints. The results of these actions are locally archived and visualized.

After that the data is transmitted to the centralized Pharos Navigator server application in the cloud. At this point, the incoming big data streams are combined with other complementary (historical) data from enterprise information systems such as ERP, CRM, etc. and archived. Then, in order to forecast any potential failures, machine learning algorithms, such as anomaly detection or outlier detection are executed on the data. As a result, the anomalies detected got flagged and the equipment with top anomalies got marked.

Finally, ERP systems are informed about change in status of the affected components. Subsequently, it triggers automated maintenance actions, such as prioritize equipment maintenance tasks, assign technicians, etc. The alarming, reporting or visualizing systems help to maintain a holistic view of the facility and support decision making processes. These functionalities are supported with technologies such as mobile devices, smart glasses or augmented reality. Moreover, it is possible to immediately start some specific, fully automatized actions, such as reinstall SW component firmware, recalibrate equipment, etc.

3.3. PdM framework implementation roadmap

The following implementation roadmap of a proposed PdM framework is based on the knowledge gained by literature study as well as by analyzing similar industrial projects. The proposal provides enterprises with advises and considerations, how to introduce PdM to the existing or new smart facilities and environments. In order to successfully implement a PdM system based on big data analytics, the following steps should be fulfilled:

- Identify relevant indicators for PdM
- Provide equipment with smart sensors or actuators
- Setup Pharos Navigator by a cloud provider
- Setup Apache Edgent with an IoT Gateway
- Connect equipment with IIoT capabilities to Pharos Navigator
- Connect equipment without IIoT capabilities to Apache Edgent
- Implement data cleansing, filtering or even machine learning algorithms on Apache Edgent
- Identify predictions with real-time and higher security constraints and implement them on Apache Edgent
- Connect Apache Edgent to Pharos Navigator by using standardized industrial protocols
- Connect ERP systems to Pharos Navigator
- Implement a data archiving in the Pharos Navigator
- Implement a predictive analytics based on machine learning algorithms in Pharos Navigator
- Connect reporting and visualizing systems to Pharos Navigator
- Connect other maintenance systems to Pharos Navigator

To begin with, it is crucial to identify all data sources which are relevant for PdM. It is beneficial, if industrial facilities are already provided with smart sensors which collect data about the ongoing processes. Even better is, if such systems are connected to the internet and are managed by remote-control systems. These are important prerequisites to implement a PdM system. However, in many cases, existing manufactures, power stations, etc. work completely separated as a silo and do not contain smart sensors or actuators at all. Right at the beginning, it is necessary to provide them with smart equipment. These components can later on gather raw process data right at the source.

Then Pharos Navigator (or similar server-side application) has to be launched in centralized server. It works as a SaaS, so it out-of-the box provides connectivity, data management or analytical components.

After that Apache Edgent (as a client-side application) is being set up on distributed servers on premises. It works as a SaaS as well and provides analytical components in real-time constraints. On one hand, in order to reduce network traffic, it is beneficial to do data filtering and cleansing processes next to the source of data. On the other hand, this solution allows for fulfilling higher level of security, since process relevant data is not sent to the Internet and stays in a “demilitarized” zone.

Then equipment with IIoT capabilities needs to be connected directly to the Pharos Navigator (CC component) or to the Apache Edgent (EC component). Subsequently, it is necessary to identify all tasks, which should be executed on edge systems or in the cloud. The transmission of the data should be done by using secure cryptographic protocols, since there are important assets of enterprises. This approach can assure data integrity and confidentiality.

Finally, third party systems are to be connected to the Pharos Navigator, such as ERP system, data reporting and visualizing systems, mobile devices, etc.

4. Experiments and results

As a proof of concept and demonstration of the proposed PdM architecture, we use a simple demo “smart environment” which is a simple HVAC application. In modern facilities, advanced HVAC systems are utilized to minimize energy consumption. Moreover, these systems can detect anomalies in facility operation and automatically trigger maintenance tasks. The functioning of smart HVAC system is based on numerous sensors and actuators, which measure and control physical values such as temperature, humidity, pressure, carbon dioxide, occupancy, etc.

4.1. PdM demonstrator

The implemented “smart building” demonstrator aims to simulate the whole PdM architecture, including data collection and anomaly detection on edge side as well as visualization and reporting on the cloud side. Figure 3 shows all the steps of the PdM prototype system for the smart HVAC application as a core of “smart building” demonstrator.

Fig. 3. PdM demonstrator process model

The prototype PdM collects a continuous temperature and humidity data streams from a digital sensor in real time (at a rate of once a minute). The sensor itself is controlled with a Raspberry Pi prototyping board. For data exchange between the sensor and the board I2C protocol is used. The prototype demonstrates capabilities of the EC. According to this implementation, the machine learning algorithm for anomaly detection runs “on the edge”, i.e. right next to the source of the data. The classification problem solution is based on recorded (historical) data. It analyses the incoming temperature and humidity measurements and in case of any suspicious behavior, assigns it (flags it with) an “anomalous” class. The anomaly detection algorithm is based on Gaussian model as it is described later.

After the analysis of the data, the edge system transfers the data to IBM Watson™ IoT Platform in the cloud (CC). In spite of the tremendous functionalities offered by IBM platform, the demonstrator uses it only for data visualization and reporting purposes. The measured values of temperature and humidity and the calculated values of anomaly(ies) are visualized with a line chart. In case of anomaly detection, an email is generated and sent. The whole architecture of the demonstrator is depicted in Figure 4.

Fig. 4. Architecture of the PdM demonstrator

In the *edge computing (EC) part* of the prototype PdM, Apache Edgent provides a convenient way to deal with real-life data as data streams. The following code segment shows the way, how to poll humidity and temperature data from HYT-221 sensor and subsequently execute anomaly detection algorithm and flag anomalous data.

```
// Polling data from sensor once a minute
TStream<SensorData> sensorData = top.poll(sensor, 1, TimeUnit.MINUTES).tag("readed");

// Execute anomaly detection and flag anomalous data
TStream<SensorData> analysedSensorData = sensorData
    .modify(sd -> {
        Boolean result = ad.isOutlier(sd.toAnalysisFormat());
        sd.setAnomalous(result);
        return sd;
    }).tag("analysed");
```

In order to detect outlier temperature and humidity data instances a Gaussian model was selected. As a first step, it was necessary to collect an unlabeled dataset $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$. It was assumed that the majority of instances in the dataset are non-anomalous.

Then it is necessary to fit model to the Gaussian distribution. It's done for every attribute x_i (where $i = 1 \dots n$), and for each of them the mean μ_i and control of variance σ_i^2 are calculated according to formula expressions (1) and (2):

$$\mu_i = \frac{1}{m} \sum_{j=1}^m x_i^{(j)} \quad (1)$$

$$\sigma_i^2 = \frac{1}{m} \sum_{j=1}^m (x_i^{(j)} - \mu_i)^2 \quad (2)$$

With the help of μ_i and σ_i^2 parameters, the Gaussian distribution can be calculated according to formula (3):

$$p(x) = \prod_{i=1}^n p(x_i; \mu_i; \sigma_i^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}\right) \quad (3)$$

This helps to find data instances having lower probability and as such being most probably anomalous. After that it is necessary to find a threshold value ε in order to distinguish normal and anomalous data instances:

$y = 0, p(x) \geq \varepsilon$ – normal data
 $y = 1, p(x) < \varepsilon$ – anomalous data

The selection of the best fitting ε is done on one hand with a cross validation dataset with labelled data. On the other hand, it is necessary to calculate F_1 score according to formula (4):

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

F_1 is a metric, which helps to identify, how many data instances are selected correctly or incorrectly, where precision and recall are calculated as following:

$$Precision = \frac{true\ positives}{true\ positives + false\ positives} \quad (5)$$

$$Recall = \frac{true\ positives}{true\ positives + false\ negatives} \quad (6)$$

Here “true positives” represent anomalous data, which were correctly classified, “false positives” represent normal data, which were incorrectly classified as anomalous, and “false negatives” are anomalous data, which were incorrectly classified as normal.

The **data communication part** of the prototype provides data transmission from the edge system to the cloud system via MQTT messages. In order to do so, the publicly available Mosquitto server was used. As shown in the following code segment, Apache Edgent provides a convenient way for data publishing:

```
// Publish data to MQTT broker
MqttStreams mqtt = new MqttStreams(top, "tcp://test.mosquitto.org:1883", "pdm");
mqtt.publish(data.asString(), "hyt-221/json", QoS.FIRE_AND_FORGET, false);
```

The data payload is sent in JavaScript Object Notation (JSON) format and is presented by the following example:

```
{
  "name": "hyt-221",
  "temperature": 21.5491294,
  "humidity": 60.53827677,
  "anomalous": false
}
```

As a **cloud computing (CC) part** of the prototype, IBM Watson™ IoT Platform provides a convenient way to set up cloud based IoT applications. It provides functionalities for connectivity, data management, data visualization and a lot more. In addition, the web-based tool called Node-RED makes it possible to easily create complex logical flows by simply wiring the necessary nodes. The HVAC demonstrator takes advantage of nodes such as MQTT subscriber, JSON parser, visualization elements such as charts and nodes as well as email notification. The server-side logic of the prototype PdM application realized in Node-RED is illustrated by Figure 5.

Fig. 5. Server-side logic created in Node-RED

4.2. Experimental setup

The prototype PdM system for HVAC application is configured for tests to include the following building elements:

- Digital Humidity and Temperature Sensor Module (HYT-221)
- Single-board [edge] computer (Raspberry Pi)
- Cloud server (IBM Watson™ IoT Platform)
- MQTT broker ([Mosquitto](#))

HYT-221 digital sensor can handle temperatures ranging from 40 °C to +125 °C with accuracy of ± 0.2 °C and relative humidity between 0% and 100% with accuracy of $\pm 1.8\%$. Thanks to the integrated signal processor, it delivers physical parameters in digital form over I²C interface. The minimal response time is around 10 seconds. Typically the sensor is used in the fields of meteorology, medical devices or agriculture [12].

Raspberry Pi is a tiny and low-cost computer to be used mainly for prototyping and learning purposes. Model A has a 700 MHz single-core CPU and it is equipped with numerous peripherals, such as GPIO, which could be used for I²C communications. From software point of view of the demonstrator a Raspbian Linux operation system and some Python and JAVA libraries were utilized [13]. A local workstation, including Raspberry Pi and the connected HYT-221 sensor is shown on Figure 6.

Fig. 6. Raspberry Pi and HYT-221 sensor

Mosquitto is an open source MQTT message broker, implementing 3.1 and 3.1.1 versions of the protocol. It offers a publicly available server/broker under the following address: test.mosquitto.org. It provides possibilities for secure and unsecure communication possibilities on the following ports: 1883 (unencrypted) and 8883 (encrypted) [14].

IBM Watson™ IoT Platform is a solution designed to develop IIoT applications. It provides functionalities for device registration and data storage and visualization. Moreover, it offers connectors with many other cloud services such as Cloudant NoSQL Database, Apache Sparks, etc.

4.3. Results of experiments

After the model creation and selection of the best fitting ε , the PdM system was ready to analyse real-time data arriving from HYT-221 sensor to Raspberry Pi computer for anomaly detection. Figure 7 shows the results of application of the anomaly detection algorithm on the data collected from a real-life **HYT-221** sensor.

Experimental results as it is shown in figure 8 show, that the implemented anomaly detection algorithm provides enough power and accuracy in detecting anomalous behaviour (outlier detection) for the combination of temperature and humidity data values. In case of the HVAC demonstrator, the anomaly detection functionality has been implemented at the edge (in **Raspberry Pi**). However, only the server-side **IBM Watson™ IoT Platform** system had the capability to generate and send email notifications in case of an outlier. The data reporting and visualization functionality was realized with the help of Node-RED nodes.

Figure 8 represents temperature and humidity values with the corresponding anomalous flag, received from the edge system via MQTT messages. The graph (a) is a subset of all measured data. It displays two anomalous combinations of temperature and humidity values at times 17:35 and 17:55, correspondingly. Chart (b) shows two gauges, imitating real-world instruments for real-time visualization of the instant sensor readings of humidity and temperature.

Fig. 7. Real-time sensor data for temperature and humidity: (a) data collected and (b) data tested for outliers

Fig. 8. Visualizing real-time data with (a) line/stock chart and (b) half-circular gauges

5. Conclusion

In our research, a distributed data-driven architecture of predictive maintenance framework is proposed, investigated and tested that combines edge computing and cloud computing solutions. On one hand, EC can help to build fully distributed PdM systems, cause it allows to execute smart data analysis on premises under real-time constrains. In addition, it enhances security by preventing from sending sensitive data to the Internet. On the other hand, CC makes it possible to integrate and analyze data streams arriving from various sources. By applying standardized interfaces, the data can be accessed and collected from different (local and remote) systems. The scalability, on-demand computing facilities of the CC allow to process data in volumes and speeds never possible before. In our proof-of-concept prototype PdM application, two readily available software solutions were utilized. To demonstrate capacities of the EC, Apache Edgent was selected, and for the CC part a Pharos Navigator and IBM Watson™ IoT Platform were considered.

The special focus of the research was on predictive analytics techniques for the proposed PdM framework. The most essential issue here is the evaluation of big data algorithms, which are suitable for real-time failure recognition and prevention. The authors tested a machine learning algorithm called anomaly detection, which represents a supervised learning algorithm, capable of recognizing anomalous behavior in measured time series data. The main objective of the proposed PdM framework was to support maintenance management decision processes with a smart maintenance tools, which are based on EC and CC and apply machine learning algorithms to solve the system state prediction problem. An implementation roadmap of the proposed architecture and a detailed description of functionalities were discussed too.

As a proof of concept, a smart HVAC system was implemented. The system consisted of a Raspberry Pi board with Apache Edgent Framework, a HYT-221 temperature and humidity sensor for EC part of the system. On the EC part it utilized the advantages of Apache Edgent framework, which allowed perceiving sensor data as data streams. In order to demonstrate the possibilities of the EC, the anomaly detection algorithm was implemented right at the edge. Usually, devices at the edge mostly execute data cleansing and filtering tasks. However, the authors believe, that ever increasing computational power of smart metering technologies and IoT gateways enables deployment of even more sophisticated data analysis algorithms at the edge. After training the model with a test dataset, the application was able to detect anomalous behavior of temperature and humidity data pairs in real-time and without the need of sending the data to the cloud. After that the processed and flagged data were sent over MQTT communication protocol to the IBM Watson™ IoT Platform for visualization and reporting purposes. For anomalies, a messages were generated and sent via email.

As a future work or enhancement to this research, we plan to explore communication options for the shop floor data with enterprise information systems. Other area of promising R&D is an innovative data reporting and visualization tools that could dramatically improve decision making processes and the daily work of maintenance personnel. The alarming, reporting and/or visualization systems help to maintain a holistic view of the smart environment and support maintenance management decision making processes. These functionalities are supported with technologies such as virtual reality and augmented reality (VR/AR) on mobile devices, smart glasses, etc. Another possible challenging task for future work would be the implementation of data-intensive predictive analytics applications in the cloud. In pursuing this task it is also important to consider, how to combine predictions, which were already executed on the edge with predictions in the cloud. This work would allow investigating the real possibilities of big data predictive analytics and the concepts of distributed dynamic data-driven application systems.

New maintenance strategies such as proactive maintenance and reliability-centered maintenance (RCM) [4] are considered as a promising maintenance strategies which not only provide smart means for failure recognition and avoidance, but also try to improve its own efficiency. In order to do so, the root cause analysis of failures is essential. It aims at identifying the source of the problems and then finding more effective maintenance solutions.

In real-life applications, to select the most appropriate management strategy, it is reasonable for enterprises to introduce risk management as well. Through risk analysis they can estimate the number and severity of the faults. The result of risk management can support decision makers in selecting the most suitable maintenance strategy or a combination of them for their industrial facilities.

6. Acknowledgments

The article was partially supported by the Academic Fund Program at the National Research University Higher School of Economics (HSE) under Academic Mobility Grant in 2018 and by the Russian Academic Excellence Project «5-100».

7. References

- [1] Predictive Maintenance 4.0. Predict the unpredictable. Report of the PwC Belgium. (2017). Available from: <https://www.pwc.be/en/documents/20171016-predictive-maintenance-4-0.pdf>, Accessed on: 2018-10-07
- [2] Williams, Z. D. (2017). IoT Analytics, In: The Top 20 Companies Enabling Predictive Maintenance: <https://iot-analytics.com/top-20-companies-enabling-predictive-maintenance>, Accessed on: 2018-03-26
- [3] eoda GmbH. (2014). Predictive Maintenance (mit R). Whitepaper, In: eoda | The Data Science Specialists: https://www.eoda.de/files/Use_Case_Seiten/Whitepaper/Predictive_Maintenance_mit_R.pdf, Accessed on: 2018-03-22
- [4] <https://cloud.google.com/solutions/oil-and-gas-equipment-monitoring-and-analytics>, (2017). Oil and Gas Equipment Monitoring and Analytics, Google Cloud Platform, Accessed on: 2018-12-22
- [5] Stamboliska, Z., Rusiński, E., & Moczko, P. (2015). Proactive Condition Monitoring of Low-Speed Machines. Springer International Publishing Switzerland 2015, 186 pages.
- [6] Gregor, M., Haluska, M., Fusko, M. & Grznar, P. (2016). Model of Intelligent Maintenance Systems, Proceedings of the 26th DAAAM International Symposium, pp.1097-1101, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-07-5, ISSN 1726-9679, Vienna, Austria DOI: 10.2507/26th.daaam.proceedings.154 Available from: <https://www.researchgate.net/publication/304657242> Model of Intelligent Maintenance Systems, Accessed on: 2018-10-07
- [7] Labib, A. (2008) Computerised maintenance management systems. In: Complex systems maintenance handbook. Springer Series in Reliability Engineering. Springer, Berlin, pp. 417-436. ISBN9781848000100. Available from: https://www.researchgate.net/publication/226606398_Computerised_Maintenance_Management_Systems, Accessed on: 2018-10-06
- [8] Patwardhan, A., Verna, A. K., & Kumar, U. (2016). A Survey on Predictive Maintenance Through Big Data. In: U. Kumar, A. Ahmadi, A. K. Verma, & P. Varde, Current Trends in Reliability, Availability, Maintainability and Safety. Springer International Publishing.
- [9] Deac, C.N., Deac, G.C., Popa, C.L., Ghinea, M. & Cotet, C.E. (2017). Using Augmented Reality in Smart Manufacturing, Proceedings of the 28th DAAAM International Symposium, pp.0727-0732, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-11-2, ISSN 1726-9679, Vienna, Austria DOI: 10.2507/28th.daaam.proceedings.102
- [10] Groba, C., Cech, S., Rosenthal, F., & Gössling, A. (2007). Architecture of a Predictive Maintenance Framework. Computer Information Systems and Industrial Management Applications.
- [11] Jeppesen, B. G. (2015). Realizing the Fourth Industrial Revolution. In: Mjolner informatics: <https://mjolner.dk/tech/realizing-fourth-industrial-revolution/>, Accessed on: 2018-10-04
- [12] https://www.ist-ag.com/sites/default/files/AHHYTM_E.pdf (2015), IST-AG. Application Note, Humidity Modul, HYT. Available from: Innovative Sensor Technology, Accessed on: 2018-10-06
- [13] https://en.wikipedia.org/wiki/Raspberry_Pi (2015), Raspberry Pi. Available from: Wikipedia, Accessed on: 2018-10-06
- [14] <http://test.mosquitto.org/> (2015), MQTT. Available from Mosquitto, Accessed on: 2018-10-06