

INDUSTRIAL GATEWAY FOR DATA ACQUISITION AND REMOTE CONTROL

Tomáš LOJKA, Marek BUNDZEL, Iveta ZOLOTOVÁ

Department of of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, tel. +421 55 602 2749, e-mail: tomas.lojka@tuke.sk,
marek.bundzel@tuke.sk, iveta.zolotova@tuke.sk

ABSTRACT

New industrial trends, related problems and proposed solutions are discussed here. The amount of data produced by various devices increases continually causing problems with remote control. We intend to address these problems with a proposed architecture, context-based services and new algorithms. We have analyzed suitable machine learning algorithms for data acquisition and selected K nearest neighbors for Event detection and Query Processing to manage data in a network of sensors and actuators. The implemented method comprises a part of the designed gateway architecture based on Internet of Things (IoT) and Serviced Oriented Architecture (SOA). We aim to improve remote control and data acquisition, monitoring and understanding of processes in technological layer in HMI interaction.

Keywords: Industry 4.0, Machine-to-Machine, remote control, SCADA/HMI, service oriented architecture, sensor network.

1. INTRODUCTION

IT is one of the fastest emerging fields merging with automation and Internet of Things (IoT) and leading to a new industrial revolution called Industry 4.0. IoT allows better monitoring and control. The principle of industrial IoT describes the Rockwell's president and chief operating officer, Don H. Davis, Jr.:

"The driving force behind productivity today isn't working faster, or working cheaper. It's working together".

Industry 4.0 expects IoT to improve the industry in terms of safety, faster interactive interoperability, improving the connection between the technological layer and the business layer, better quality of monitoring, new concepts of device and data management over many more devices and much larger area. We address here the management of many devices to improve data collection and integration, control, data storage and visualization.

Considering the needs and the problems we describe here our proposed architecture based on SOA and IoT. We focus on the implementation of data management, energy consumption and network overload reduction, data integration and context based services. The proposed architecture offers services through a unified connection to HMI (Human-Machine Interface) and MES (Manufacturing Execution System).

The architecture is implemented partially. We describe here the sensory data management. Our intention is to collect data from the sensors and communicate it to the actuators via the content or the context of the messages. This requires a fixed form of communication. The sensors and actuators network lacks the self-configuration capability. Therefore, we analyze data mining techniques and WSN to be used in a solution for an adaptive gateway context configuration. The other main contribution of this paper is development of a context-based service in the gateway. This service offers data to clients like HMI based on the context or the area of the interest. This will save time selecting the tags to be sent to the clients (hundreds of tags are used in the industry). The user does not need to perfectly understand the technological processes to select the tags to be monitored. The user knows the area or the

topic of interest and the rest handles the self-management of our gateway. We focus on data collection and its representation to HMI clients mainly. This forms a part of the remote control, where has big potential a wireless communication [1].

This paper describes also the remote control, the future of SCADA/HMI and our concept and the realized gateway functionality with the experiments. We describe the remote control and the proposal of the future SCADA/HMI solution that influences the remote control greatly.

2. REMOTE CONTROL WITH HMI

The remote control enables to control the technological process from a remote location. The remote control has been developed from the local control functionalities and implemented with a wider access to it. This wider access to SCADA enables the access to the monitoring and control. Remote controlling is traditionally done via the Internet. Three main architectures using the Internet for the remote control exist [2]:

- One-to-one - is based on a direct connection between the user and the machine.
- Many-to-one – Enables multiple users to connect to the controlled machine.
- Many-to-many – Enables multiple users to connect to multiple controlled machines. In many-to-many architecture SOA is applied. There are solutions based on web services, Jabber/XMPP, SOCRADES, ubiHome and others. These architectures do not implement a description of the controlled environment nor a description of the machine's structure. Therefore, we want to create a solution operating with a description of the controlled environment.

Our intention is to create a network not only producing information but also offering data usable to improve the description of the monitored environment, following the latest trends in the industry. We also describe some features of the future SCADA/HMI.

3. FUTURE SCADA/HMI

Industry 4.0 brings new possibilities into communication. Communication influences the current SCADA system model in terms of data acquisition. Let us consider a traditional plant producing iron plates. Usually, it features a centralized and a distributed technology. The distributed part consists of Programmable Logic Controllers (PLCs), protection relays, sensors, fault components, a remote substation, and an industrial network gateway. SCADA represents the centralized part and it is connected to OT (Operation Technologies) parts. SCADA collects and transmits data such as pressures, distances, velocities, temperatures, regulation data and other from the distributed part of OT (or from Industrial Automation). SCADA is then used for data presentation and visualization, supervisor control and for data analysis (SDF – Statistical Data Processing). An example of SCADA visualization of a calcic factory is in Fig. 1.

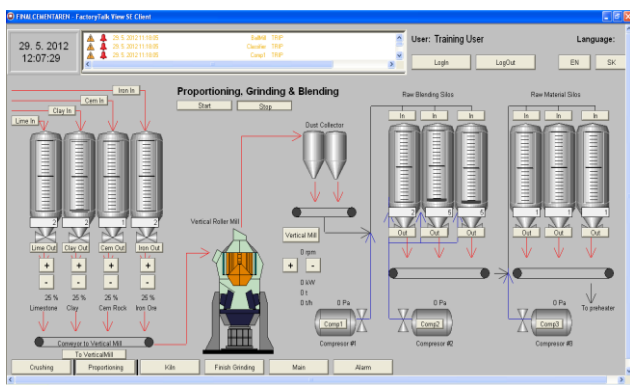


Fig. 1 Example of a calcium factory

SCADA coupled with IoT may reform in the future. Data acquisition may change in the context of IoT to improve the remote control. Many companies produce sensors and devices offering intelligent functionalities enabling decentralized intelligence and decision taking. Machines communicate with machines (M2M) and decrease human intervention. Data can directly enter the Enterprise Resource Planning (ERP) system by M2M and the intelligent functionality of sensors or machines. SCADA may be spread into a distributed system but still maintains its centralized role thus improving safety, security and efficiency of the whole plant [3].

At present, SCADA is integrated into productions generating large amounts of data from the control systems, the machines and the sensors. Data are integrated in one place and after that offered to MES and ERP systems. This process will change in the new architecture [4]. Every "thing" will be capable to communicate and cooperate [5]. Via the improved communication the people and the system can make better decisions. For example: NEST thermostat is a device in Level 0. There is implemented an intelligent industrial model that allows this device to work unsupervised.

The new industry consists of end-services which create a big population of intelligence devices. This high granularity architecture is concentrating to apply intelligence on distributed nodes. The traditional applications concentrate intelligence on a few large

applications only. The characteristics of the new SCADA in the context of IoT explain these points:

- Better representation of data in HMI thanks to the pervasive computing - distributed computing with distributed intelligence on the devices.
- M2M interaction – the devices can communicate with each other and with the help of distributed intelligence can solve quickly, safely and unsupervised almost every situation.
- Better safety and availability due the intelligence and self-management of devices

The fact that devices can communicate with each other or with humans and use the distributed machine intelligence will change the present SCADA system model. The requirements of improved flexibility and adaptation motivate the "upgrade" of the present SCADA system model that is in use for almost 25 years [6].

4. CONCEPT DEFINITION OF DATA ACQUISITION AND MANAGEMENT IN OUR SOLUTION

We have proposed a complete architecture and now we work on a method to fulfill the criteria defined for the proposed gateway [7].

We defined a concept consisting of 3 layers (see Fig. 2):

- Technological processes
- Data processing and routing
- HMI, MES, database.

Technological processes – we only focus on CPS in this layer. This layer is almost finished. We have implemented algorithms reducing the communication and energy consumption in the CPS system. The communication reduction is important to reduce the network overload in a large network of devices. We designed/modified and applied online non-cyclic algorithms inspired by data mining techniques into CPS. We use a classification algorithm to determine the importance of a message for the created topic in MQTT protocol [8].

Gateway – We have designed a gateway that enables the access to the data from the CPS systems. The gateway integrates OT and IT. The gateway offers a unified connection to the remote control and to the MES system. With the implemented SOA and the context-based service data are accessible more conveniently and comprehensibly. We have proposed the improvements for the remote control and MES. We have designed the runtime and the management parts of the gateway. These parts are subject to the research aimed to improve or to design the machine learning algorithms. They help to control the acquired data and offer a unique view on the technological processes without the knowledge of what is behind the gateway. The gateway plays the key role in SCADA. Therefore, the designed gateway is peer-to-peer connected with other gateways to provide stable functionalities.

HMI, MES and the database – these are the clients of our gateway. The gateway offers data from the technological processes either as a response to a query or in

the form of higher abstracted data context created by our gateway. This will enable to improve the human-machine interaction and understanding of states and situations in the technological processes.

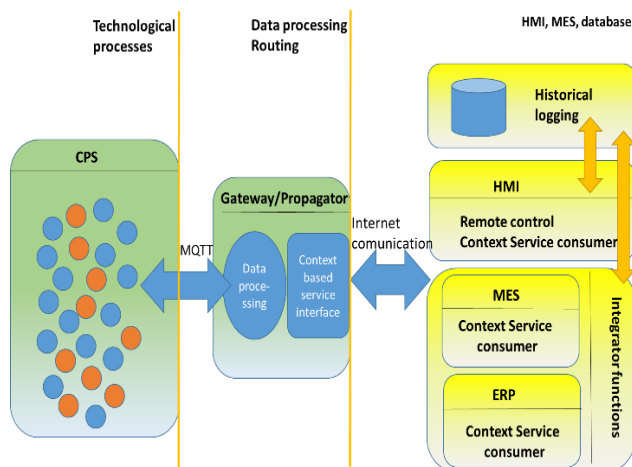


Fig. 2 Symbolic schema of solution

4.1. Machine learning

We have implemented machine learning (ML) to perform data management and self-configuration in the gateway. We have identified the methods of the potential use. The aim is to implement the data management and the management of devices from the IoT gateway to create an abstract layer of the sensors and actuators network.

ML also addresses the analytical decision making that uses data mining algorithms to discover information and to retrieve knowledge from the data. ML may help to retrieve information obscured from the human operators by the sheer number of the variables relations frequent small changes of the values etc. The industrial IT environment is suitable for ML implementation (regarding the storage, knowledge and computing). ML is not a new element in industry. It has been implemented in MES, ERP or HMI systems but not much on the low levels of the devices data management.

We have focused on the time-oriented data from the devices and we plan to implement ML in the proposed gateway in our work. We aim to use ML to create distributed data processing. The processed data will be then offered to the regular clients, the cloud, or other system like ERP or MES via a unified connection.

ML encompasses techniques for [9]:

- Retrieving knowledge from new data
- Making predictions about new data from the old data

The typical ML analytical applications are Classification, Clustering, Pattern discovery and Prediction. We look at these methods as on functional applications of ML in a network of devices.

4.2. Functional applications of Machine Learning

Small or mobile network device solves a problem of its memory and power's constrains. The devices usually use wireless network. Therefore, the devices network is similar to a Wireless Sensor Network (WSN). WSNs' are subject to intensive research [10]. We have selected methods related to WSN described in [11] that are divided into five functional fields:

- Routing in WSNs – the goal is to find an optimal path by learning from experience. Routing determines the optimal routing paths to reduce the cost, increase the efficiency and to achieve real time routing and the best performance and quality of services.
- Clustering and data aggregation – suitable for large networks, where it is not effective to transmit data directly to the sinks [10]. Data are first aggregated and then transferred to the sinks.
- Event Detection and Query Processing – the fundamental methods for any large network of devices.
- Localization and object targeting – ML is used to identify positions of the devices without using GPS.
- Medium Access Control (MAC) – a protocol that is designed for effective data transmission in WSN. MAC manages energy consumption and latency. It helps to identify the duty cycle in the sensor network.

The non-functional aspects are not covered here. We have selected Event Detection and Query Processing functionality group for the final system implementation. We want the gateway to manage the sensors and the actuators. The Event Detection and Query Processing functionality group is the most suitable for it.

4.3. Event Detection and Query Processing

This method reduces human intervention what is the goal of M2M. ML serves to develop effective event detection mechanisms spending less computing and storage resources in. In the case of a query, ML helps to find locations of the devices that are interesting to monitor and to control. Therefore, the message can be sent to a selected group of devices only.

This functional group contains algorithms that are under research (mainly the algorithms for WSN). The simplest algorithms are based on thresholding, but simple thresholding is not suitable for all situations. The goal is to use ML to design and to develop an intelligent event detection and query system. The algorithms of this functional group are described in [11].

The monitoring in the case of event detection and query processing can be divided to event-driven, query-driven and continuous [11].

Table 1 ML-based detection and query processing solutions variants [11].

| Approach | ML | Data delivery models | Complexity | Characteristics |
|---|----------|----------------------|--------------|---|
| Event region detection | Bayesian | Event-driven | Low Moderate | Fault-tolerant event region detection Real-time activity recognition |
| Activity recognition | k-NN | Query-driven | Low Moderate | In-network query processing |
| In-network query processing Query processing in 3D space | NNs | Event-driven | Moderate | Real-time and lightweight forest fire detection |
| Forest fire detection | DT | Event-driven | Low | Disaster prevention system |
| Query optimization | PCA | Query-driven | High | Query optimization and dimensionality reduction |

The *k*-means clustering was the potential method to be used in our implementation. This method is suitable for query processing [12].

We used the *k*-means clustering method to build the mapping tables. To do that we proposed to build a self-management layer of sensors and actuators. The publish/subscribe method plays a key role here.

4.4. The mapping tables

The publish/subscribe method that we have used is also known as pub-sub. It is a paradigm for asynchronous communication between the entities. The subscribers are the nodes that need to receive the event messages from the publishers to which they are subscribed. The publishers are the nodes creating events and sending the messages to all the listeners [12]. Publish/Subscribe has a big potential in distributed applications for data acquisition systems. Publish/Subscribe is based on context-based subscriptions. The problem was to identify which subscriber from the mapping table should be associated with a specific context group or groups.

4.5. The gateway and the experiments

We have developed the proposed gateway in MS Visual Studio. We have chosen C# to speed up the development and verification. The gateway consists of:

- **The logic part** – is the main part of our gateway. The gateway is a complex and we describe here the relevant parts only. We have used *k*-means clustering. The number of the neighbors in *k*-means clustering represents the number of the monitored/controlled areas in our case. We used Euclidean distance (1) as the distance measure:

$$D = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (1)$$

where: D is the distance, x_i is the sensor vector consisting of [temperature, location in line], y_i is the vector of temperature and location means.

We have used simulated sensors only for testing and verifying the gateway functionality. One of the experimental results is in Fig. 3 in the developed graphical part.

- **The graphical part** – we have created this part in WPF. We have created a class for the graphical representation of the nodes. This class contains the designed properties, the positioning behavior and the appearance of a node. The graphical view is updated according to its representation (sensor/actuator) from the logical part. The graphical part was created to demonstrate the states of sensors/actuators in network only.

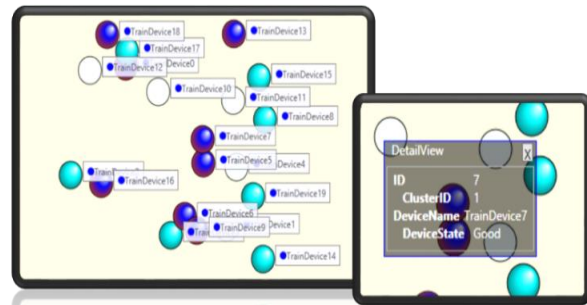


Fig. 3 Graphical representation of the gateway's context-based clustering for creating the mapping tables

As a result, we want to create a routing table. This table is created from the results of the *k*-means clustering method, which we applied on the sensors. Every cluster found by *k*-means clustering represents a topic (a context). This topic closely specifies and explains which sensors, actuators, gateways or clients (HMI, reporting applications) can be used to send or receive messages in the network.

We have used 20 sensors defined by ID, the mean temperature and the location in the line in our experiment. We have defined the position of the sensors to be integer numbers defining their location in the line. This number defines the sensor's position in the monitored area.

We have simulated the measured values of the sensors. According to these values the algorithm grouped the sensors into three areas. These areas are offered to the clients, the remote controllers for the monitored areas. In this case, our algorithm was sensitive to the position of the sensor and the measured temperature. The gateway grouped the sensors into dynamically resizing groups as they appeared in the network. The resulting mapping table is in Fig. 4.

According the mapping table, the user or the developer of the HMI/MES application can select the monitored area or even the mean of the monitored area (in the case of a large group of sensors) and to monitor it easily, to calculate trends or to control the monitored area.

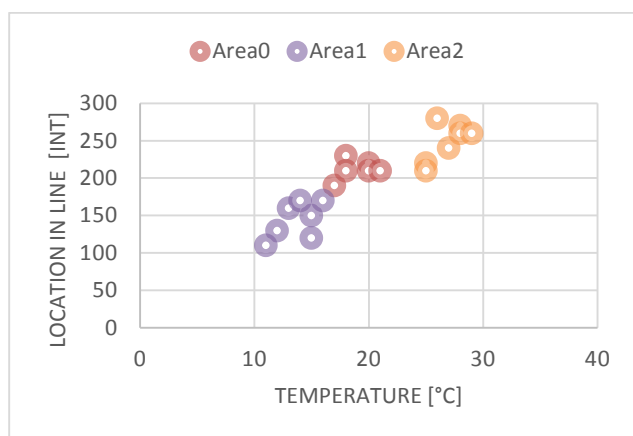


Fig. 4 Graphical representation of the communication groups automatically grouped into three monitoring areas according to their measurements.

5. FUTURE WORK

We will be focus on the development of a unified connection to the gateway to allow connections of HMI application, database, MES and Cloud services to retrieve data from gateway in the context form. We want to verify and to improve our method implemented in the gateway and the CPS systems. The verification will be based on performance testing and the verification of the HMI and MES applications result correctness. Final result of our work will to bring contribution for smart manufacturing to operate with real-time data , making intelligent decisions, predicting situations in real-time as it describes reference [13].

6. CONCLUSIONS

We have analyzed data mining techniques for data management in our gateway. We have used k -means clustering based method to create a mapping table. This table enables routing data between the sensors/nodes and the gateway using the publish/subscribe method. The contribution lies in better data acquisition and the development of the context-based interface for clients like HMI, MES or database services that will improve the remote control.

Our goal was to create an interconnection between all devices (smart products, smart machines) to reduce human intervention in not important processes. Data acquisition is still very important and SCADA functionality will be preserved but the way of data acquisition will change. Not only SCADA will have a connection to the technological layers, but also the higher business layer, which will analyze the real-time data faster. With the improved flexibility of the devices, the data will be better accessible to other systems. That is the intention of our gateway's method.

ACKNOWLEDGMENTS

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded

by the ERDF (50%) and by grant KEGA - 001TUKE-4/2015 (50%).

REFERENCES

- [1] VAŇUŠ, J. – KUČERA, P. – MARTINEK, R. – KOZIOREK, J.: Testing of the Wireless System Reliability for Technical Functions Control in the Smart Home. *Human-centric Computing and Information Sciences*. Springer, 2015, pages 2-19.
- [2] GUO, X.: Remote control service system architecture and dynamic web user interface generation, a Doctoral Thesis, 2011.
- [3] VALE, Z., et al.: Distributed energy resources management with cyber-physical SCADA in the context of future smart grids." *MELECON 2010-2010 15th IEEE Mediterranean Electrotechnical Conference*. IEEE, 2010.
- [4] PENIAK, P. – FRANEKOVÁ, M: Trendy v integrácii riadiacich a informačných systémov v podmienkach výrobných podnikov. *ATP Journal* 03/2012. ISSN 1335- 2237, pp. 32- 33.
- [5] KRANENBURG, R. van – ANZELMO, E. – BASSI, A. – CAPRIO, D. – DODSON, S. – RATIO, M.: The internet of things, *Proc. 1st Berlin Symp. Internet Soc.*, Berlin, Germany, 2011, pp. 25–27.
- [6] BLOEM, J. – DOORN, M. van – DUIVESTEN, S.– EXCOFFIER, D. – MAAS, R. – OMMEREN, E. van: *The Fourth Industrial Revolution Things to Tighten the Link Between IT and OT*. Sogeti VINT 2014.
- [7] ZOLOTOVÁ, I. – BUNDZEL, M. – LOJKA, T.: Industry IoT Gateway for cloud connectivity, *Advances in Production Management Systems: Innovative Production Management Towards Sustainable Growth*. Springer International Publishing, 2015, pp. 59-66.
- [8] ZOLOTOVÁ, I. – LOJKA, T.: Online data stream mining in distributed sensor network, *WSEAS Transactions on Circuits and Systems*. Vol. 13 (2014), pp. 412-421. - ISSN 1109-2734.
- [9] YASAKETHU, S. L. P. – JIANG, J.: Intrusion Detection via Machine Learning for SCADA System Protection, *Proceedings of the 1st International Symposium for ICS & SCADA Cyber Security Research*, 2013.
- [10] LOJKA, T.: Sensor networks autonomy, *SCYR 2014: 14th Scientific Conference of Young Researchers : proceedings from conference*, May 20th, 2014, Herľany, Slovakia. - Košice : TU, 2014 S. 239-242., ISBN 978-80-553-1714-4.
- [11] XIONG, N. – SVENSSON, P.: Multi-sensor management for information fusion: issues and approaches. *Information fusion*, 3(2), pp. 163-186, 2002.
- [12] ALSHEIKH, M. A., et al.: Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *Communications Surveys & Tutorials, IEEE* , vol.16, no.4, pp.1996-2018, 2014.

- [13] KUMARAGURU, S. – KULVATUNYOU, B. – MORRIS, K. C.: Integrating Real-Time Analytics and Continuous Performance Management in Smart Manufacturing Systems. *Advances in Production Management Systems. Innovative and Knowledge-Based Production Management in a Global-Local World*. Springer Berlin Heidelberg, 2014. 175-182.

scientific research is focusing on data acquisition and HMI system.

Marek Bundzel was born 8.5.1974. He holds a PhD. degree in artificial intelligence. Presently he is the head of the Center for Applied Cybernetics at the Department of Cybernetics and Artificial Intelligence at the Technical University in Košice.

Received July 7, 2015, accepted August 25, 2015

BIOGRAPHIES

Tomáš Lojka was born on 9.10.1988. He graduated (MSc) from the Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics at Technical University in Košice. Now he is PhD student at the Department of of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics at Technical University in Košice. His

Iveta Zolotová graduated (MSc) from the Department of Technical Cybernetics of the Faculty of Electrical Engineering at Technical University in Košice. She defended her CSc. in the field of hierarchical representation of digital image. Since 2010 she has been working as a professor at the Department of Cybernetics and Artificial Intelligence at the Faculty of Electrical Engineering and Informatics at Technical University in Košice. Her scientific research is focusing on SCADA/HMI, Internet of Things and web labs. In addition, she also investigates questions related with the digital image processing.