IMEKO TC6 International Conference on Metrology and Digital Transformation
September 19 – September 21, 2022, Berlin, Germany

# THE GENERAL APPROACH OF SENSOR NETWORK

*Marek Havlíček* [a], *Martin Koval* [a, *], *Jiří Tesař* [a]

[a] Czech Metrology Institute, Brno, Czech Republic, email address: mhavlicek@cmi.cz
* Corresponding author. E-mail address: mkoval@cmi.cz

***Abstract*** – The Sensor Network (SN) [1] has become already a well-established concept widely used in many areas of industry such as automotive, mechanical engineering, food industry and energy supply, health care, etc. Rapid development of new technologies has brought novel approaches and implementation of new technologies such as Artificial Intelligence (AI), Cloud Computing, Big Data storage, management, etc. Current trends focus predominantly on monitoring and optimization of processes in order to increase effectiveness and reliability which can be effectively reached only with the help of sensor networks. This article focuses on the general sensor network structure, chances and pitfalls which are related to their use.

***Keywords***: Sensor Network, Uncertainty Measurement, Artificial Intelligence, AI.

## 1. WHAT IS THE SENSOR NETWORK LIKE?

The sensor network can be understood as a backbone of the processes which need to be monitored and/or optimized. In the system where many different sensors provide a complex overview of the ongoing processes, a model representing such an ensemble can be created. We can refer to this model as to a Digital Twin [2] of the system. The Digital Twin enables a real-time system monitoring, and processes history analysis and effective prediction of future events which can be forecasted with the use of advanced algorithms and AI. The result of such an interplay between the sensor network and the effective feedback control is the minimization of negative events within the system.

## 2. SENSOR NETWORK STRUCTURE

The basis of the Sensor Network comprises of the Input, Signal Processing and Output. These three areas can be further extended according to their specific use in particular processes. The basic type of the SN consists of sensors of the same type with fixed network topology. These sensors usually send data to the Data Processing Unit in fixed intervals. The output may consist of the processed measured data with metadata which store additional information about the process history. An example of such process could be a system for temperature monitoring of sensitive goods according to BS EN 12830:2018.

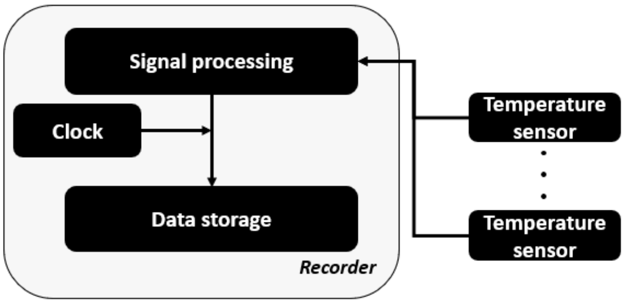


Fig. 1. An example of a Simple Sensor Network according to EN 12830:2018.

Many more factors have to be taken into account in complex systems, e.g., dynamic topology of the network, Big Data, different sensor types, location of data processing, complex mathematical algorithms, prediction, security, etc. It will be discussed in more detail in the following chapters.

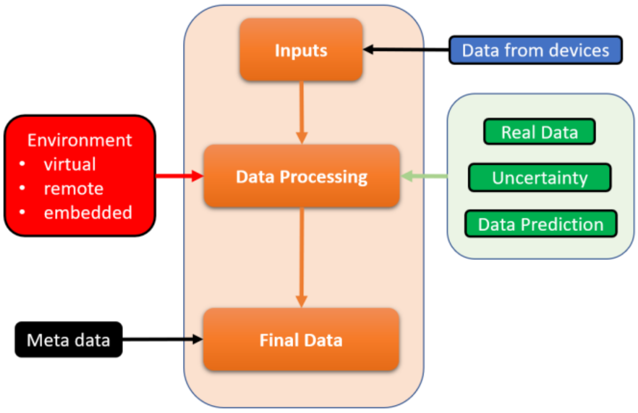


Fig. 2. An example of Complex Sensor Network.

### *2.1 Sensor Network Inputs*

The Sensor Network may consist of many devices/sensors which can widely differ in numbers, complexity, signal types and data formats. Units of simple sensors as well as thousands of sophisticated devices can both form a structure which can be considered as the Sensor Network. Considering the amount of data collected and processed in such network, we may face challenges with their processing as the amount of data increases.

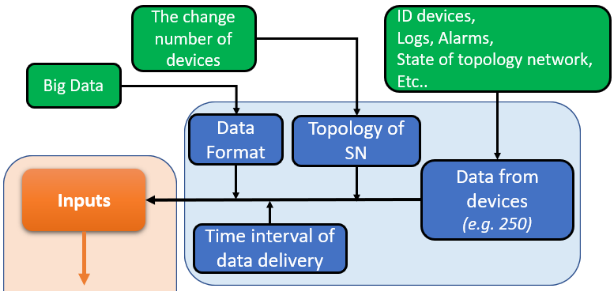

The change number of devices
ID devices, Logs, Alarms, State of topology network, Etc..
Big Data
Data Format
Topology of SN
Data from devices (e.g. 250)
Inputs
Time interval of data delivery

Fig. 3. An Example of a Complex Sensor Network – Inputs.

In such case we talk about Big Data. The Big Data can be well described as the 5V model: Value, Volume, Veracity, Variety and Velocity. Each "V" has its specific influence on the SN architecture [3].

The Value represents the usefulness of the data. In the SN data hierarchy, the data priority is set from the most to the least important. There is an evident difference in the value of the real measured data, which are used for calculations, and the metadata of measurement data. The Value provides very useful information which can help to interpret data more accurately. One important group of information comes from, e.g., alarms. The alarms also have their own hierarchy which defines their roles in informing about the limits of sensors and correct functioning.

The Volume of the data generated in the SN depends on the recording frequency and the number of variables which are recorded. If just the measured data with the corresponding metadata are stored, the amount of such data can reach typically TB or PB levels. In the case that the complete data sets including text and graphics are transferred, the volume of the data can overcome EB levels and can go even beyond that.

The Veracity represents the quality of information, their uncertainty or accuracy. The information can be inherently inconsistent, non-complete, ambiguous or its reliability can be reduced. These facts form a set of requirements which have to be applied so that it can be decided which data can be used for further analysis.

The Variety in the SN describes different forms of information. The data coming from various types of sensors and appliances can be transferred in either structured or unstructured form. In the first case, the subsequent separation and analysis is relatively easy. The situation is diametrically different for unstructured data. In such case, the data mining, sorting and analysis is more complex which may result in errors in extracted data sets.

The Velocity in the SN is the key parameter which describes the speed of the data transfer and processing. Combinations of various sensors and data structures influence the final data transfer and processing velocity which correlate with the computational power needed. In some applications, a real-time process monitoring is necessary such as in the medical applications or nuclear power plants, where any delay may have fatal consequences.

One of the important aspects which play a crucial role in Inputs is the configuration of the Sensor Network topology. Different time of data delivery from distributed sensors has to be also considered. Another important factor for Inputs, which play an important role, are the dynamic changes of participating sensors. The sensors may be disabled, replaced, maintained, damaged or exposed to disturbances which can possibly have a significant effect on the whole Sensor Network.

## *2.2 Data Processing*

Data Processing can be considered as the core of the Sensor Network. This task can be divided into separate fields which need an individual approach. The Data Processing can include the real measured data as well as the predicted data with their corresponding uncertainty. The data prediction has already become an integral part of the state-of-the-art Sensor Networks.

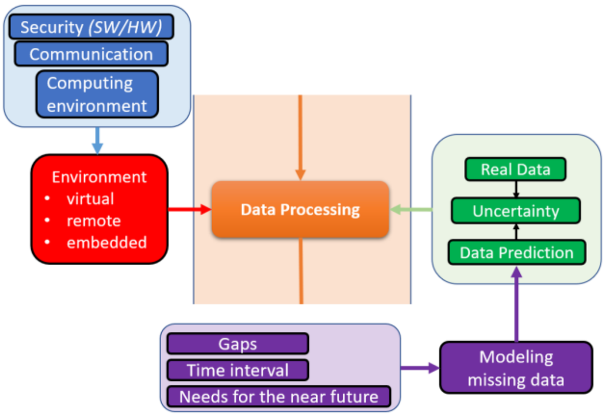


Fig. 4. An example of the Complex Sensor Network – Data Processing.

### *Data Prediction*

The effectivity and reliability of the data prediction is crucial for the modelling of specific missing or corrupted data. Reliable data prediction on different timescales (minutes, hours, days, etc.) and the information about their uncertainties are necessary. Another aspect which plays an important role is the communication loss due to various reasons (service, calibration, etc.) or due to the failure of sensors. Another application of the Sensor Network is the prediction of the network topology in the near future based on historical data. A careful data analysis may help to predict the situations which will occur during the expected events and prepare adequate measures to cope with them, such as maintenance, overloads etc. These data can be modelled with the use of various algorithms based on the Artificial Intelligence (AI). Nowadays, the progress in the AI development is accelerating. It mainly focuses on three areas which can be characterized as computational learning, reasoning and self-corrections. All these aspects can be directly applied in the SNs. The machine learning focusses on the data mining and creating rules for their conversion into useful information. The machine reasoning aims at searching for the most convenient algorithm from the family of available solutions and its implementation in the particular process. Automated self-correction mechanisms are employed in many processes in order to reach the best results in particular processes. The AI can be divided into different categories based on its capabilities: Artificial Narrow

Intelligence (ANI), Artificial General Intelligence (AGI) an Artificial Super Intelligence (ASI). The ANI is frequently used in different applications, the AGI and the ASI are still subjects of research. Another categorization of the AI into four classes is based on its functionality (see Table 1) [4, 5].

Table 1. AI basic categories overview [4, 5].

| Type AI | Use |
|---|---|
| Reactive AI<br>(type ANI) | effective for simple classification and pattern recognition tasks;<br>incapable of analyzing scenarios that include imperfect information or require historical understanding; |
| Limited memory<br>(type ANI) | can handle complex classification tasks and use historical data to make predictions;<br>capable of completing complex tasks (e.g., autonomous driving);<br>needs big amounts of training data to learn tasks;<br>vulnerable to outliers or adversarial examples; |
| Theory of mind<br>(type AGI) | should be able to provide results based on an individual's motives and needs;<br>training process would have a lower number of examples than type ANI; |
| Self-aware AI<br>(type ASI) | should be aware of the mental state of others entities and itself;<br>it is expected to outperform human intelligence; |

### *Environment*

Another important factor for the Data Processing represents the environment where the data are physically processed. Current technology enables the use of a variety of different virtual environments such as Cloud Computing, remote servers or special-purpose built-in computers. The real location of the data processing influences also the quality of the Sensor Networks.

In the case of virtual environments, the problem with insufficient power is not necessarily the limiting factor. One of the most important parts of the SN is the cyber security of the environment, communication channels and sensors themselves. From the security point of view, the SN begins at sensors. If data reliability shall be guaranteed then all sensors have to be secured from the HW and SW point of view. The HW security is essential in order to prevent any unauthorized change of parts containing the SW, which could possibly compromise the measured data. The HW security can be realized in a non-destructive or destructive way. Any unauthorized access to the sensor/device results in its destruction in the case of a destructive solution [6]. Any fraudulent data manipulation using such damaged sensor is either physically impossible or technically challenging. The non-destructive solutions typically involve different ways of sealing, which indicate an unauthorized access into HW parts. It is worth noting that the availability of technologies which are capable to substitute HW parts is higher than in the past.

In the case that sensors contain SW it is necessary to deal with the security from the SW point of view which shall include a basic minimum of the integrity check, authenticity and alarms. In the case of more advanced sensors with bi-directional communication where the remote control of sensors is possible, calibration parameters are available, etc., it is essential to secure access rights. If the system parameters can be changed, it is a good practice to use an event logger in order to guarantee the traceability of changes. One of the important factors which help with the data analysis is the presence of the metadata related to the particular data file. This metadata contains additional information which may be essential for a subsequent analysis. One of the most effective ways for the metadata protection represents a blockchain list of records [7]. For the network itself, communication is an essential prerequisite. Hundreds of different communication solutions including protocols and interfaces are now available on the market. Criteria which shall be considered during the SN communication design include energy demand, network type, compatibility, security, open source, etc. The SN design shall also include a risk analysis [8].

### *Uncertainty Evaluation Methods*

Uncertainties in the field of Sensor Networks represent a crucial aspect which should be always taken into account. Each sensor should be considered as an independent device placed in a certain environment and it should be treated as such.

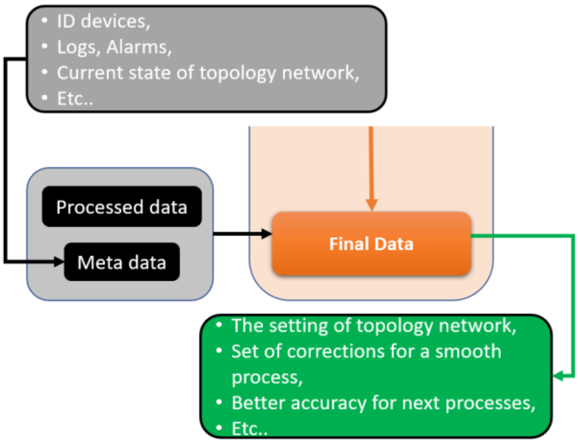
The uncertainty evaluation in the field of metrology is the integral part of all processes where the measurement is realized. Depending on the processes and the field of measurement, the used models can vary substantially. In the case of SNs, the uncertainty evaluation can be challenging. Relatively simple SNs working with basic measurement models and consisting of a few types of sensors represent the case in which the standard procedures can be applied. In the case that the data are summarized in regular intervals and only one process is monitored, then it is possible to use The Law of Propagation of Uncertainties (LPU) [9], Monte Carlo [10], etc. In the case of more sophisticated SNs, it is necessary to use mathematical models which are suitable for the particular situation. In the case that some data are not available at the moment or were removed from the data set models like Bayesian statistical models [11], Fuzzy theory, etc., should be used. An overview of commonly used methods for the uncertainty evaluation is shown in Table 3.

Table 3. Example of using Uncertainty evaluation methods [3,8-11].

| Type Size | Use |
|---|---|
| LPU | general uncertainty evaluation for complete datasets; |
| Monte Carlo | uncertainty evaluation for asymmetric and inadequate datasets; |
| Shannon’s Entropy | determining the amount of missing information on average in a random source; |
| Fuzziness/ Fuzzy Theory | processing of vague or ambiguous datasets for complex models; |
| Bayesian Statistical Models | they are particularly useful when there exists information about the true value of the measurand prior to obtaining the results of a new measurement; |

### *2.3 Sensor Network Output*

The output of the SN depends on the particular application. Examples shown in Figures 1-5 imply that the output can consist not only of the measured data but also contains the metadata which can enable more efficient data processing. The metadata can contain the data directly recorded by the sensors or they can be generated during the data processing (e.g., actual topology of the SN), alarm analysis, event logger messages, sensor IDs, etc. Further utilization of the data depends on the particular application. The outputs can be used in different ways such as process indicators, triggers (process breaks, notifications) or for analyses. Alternatively, the analysis can be directly in a machine-readable format which can be directly used by another SN with minimal changes in the configuration.



• ID devices,
• Logs, Alarms,
• Current state of topology network,
• Etc..
Processed data
Meta data
Final Data
• The setting of topology network,
• Set of corrections for a smooth process,
• Better accuracy for next processes,
• Etc..

Fig. 5. An example of the Complex Sensor Network – Final Data.

## 3. CONCLUSIONS

The SN is being implemented into many processes around us. The SN is foreseen as a part of smart cities, smart grids, complex processes monitoring in industry, autonomous driving, medicine and many other applications. Together with other technologies such as the AI and with the utilization of the Big Data, the SN is becoming an important tool for effectivity optimization of many processes. The SN has helped to push the limits in metrology towards new effective algorithms in the AI or in challenges in the uncertainty evaluation related to the utilization of the AI as well as in the implementation of solutions for digital transformation.

## ACKNOWLEDGMENTS

This work was funded by the Institutional Subsidy for Long-Term Conceptual Development of a Research Organization granted to the Czech Metrology Institute by the Ministry of Industry and Trade. Project number: UTR 22E601104.

## REFERENCES

[1] Hairong Qi, S. Sitharama Iyengar, Krishnendu Chakrabarty. *Distributed sensor networks - a review of recent research* Journal of the Franklin Institute, 2001. DOI: 10.1016/S0016-0032(01)00026-6.

[2] Andreas Deuter, Florian Pethig, *The Digital Twin Theory*, Project: Asset Administration Shell (Industry 4.0), Munich, 2019. DOI: 10.30844/I40M_19-1_S27-30.

[3] Reihaneh H. Hariri, Erik M. Fredericks and Kate M. Bowers, *Uncertainty in big data analytics: survey, opportunities, and challenges*, Journal of Big Data, 2019, DOI: 10.1186/s40537-019-0206-3

[4] Hanif Khan, *Types of AI | Different Types of Artificial Intelligence Systems,2021*, https://www.researchgate.net/publication/355021812

[5] Linda Tucci, *A guide to artificial intelligence in the enterprise* Reihaneh H. Hariri, Erik M. Fredericks and Kate M. Bowers, *Uncertainty in big data analytics: survey, opportunities, and challenges*, Tech Target-Search Enterprise AI: E-Guide, 2021, https://www.techtarget.com/

[6] Nekoogar, Faranak, Dowla, Farid, Twogood, Richard, Lefton, Scott, *Secure RFID tag or sensor with self-destruction mechanism upon tampering*, United States Patent: Document ID: US 20150339568 A1, 2016.

[7] D. Peters, J. Wetzlich, F. Thiel and J. -P. Seifert, "*Blockchain applications for legal metrology*," 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2018, DOI: 10.1109/I2MTC.2018.8409668.

[8] ISO/IEC, "ISO/IEC 27005:2018 Information technology - Security techniques - Information security risk management," International Organization for Standardization, June 2011.

[9] LUM, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, JCGM 100:2008, GUM 1995 with minor corrections. BIPM, 2008.

[10] Monte Carlo, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement' – Propagation of distributions using a Monte Carlo method, JCGM 101:2008. BIPM, 2008.

[11] Bayesian Statistical Models, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models, JCGM GUM-6:2020. BIPM, 2020