Automation, Robotics & Communications for Industry 4.0

Proceedings

of the 1st IFSA Winter Conference on Automation, Robotics & Communications for Industry 4.0 (ARCI' 2021)

3-5 February 2021, Chamonix-Mont-Blanc, France

Automation, Robotics & Communications for Industry 4.0

Proceedings of the 1st IFSA Winter Conference on Automation, Robotics & Communications for Industry 4.0 (ARCI' 2021)

3-5 February 2021

Chamonix-Mont-Blanc, France

Edited by Sergey Y. Yurish

IFSA Publishing

Sergey Y. Yurish, *Editor*
Automation, Robotics & Communications for Industry 4.0
ARCI' 2021 Conference Proceedings

E-mail (for orders and customer service enquires): ifsa.books@sensorsportal.com

Visit our Home Page on http://www.sensorsportal.com

ISSN: 2938-4788
ISBN: 978-84-09-27538-0
BN-20210129-XX
BIC: TJFM

Contents

Foreword

On behalf of the ARCI' 2021 Organizing Committee, I introduce with pleasure these proceedings with contributions from the *1st IFSA Winter Conference on Automation, Robotics & Communications for Industry 4.0 (ARCI' 2021)*, 3-5 February 2021.

According to the modern market study, the global Industry 4.0 market will reach US$ 155.30 Billion by 2024 growing at the CAGR of slightly above 14.9 % between 2018 and 2024. Increasing adoption of the industrial internet and IIoT worldwide in manufacturing units, growing focus on enhanced efficiency of machinery and systems, and reduced production costs play a significant role in the growth of the market worldwide.

Industry 4.0 represents the 4th industrial revolution that marks the rising of new digital industry. It is defined as an integrated system that comprises numerous technologies such as advanced robotics control, automation tools, sensors, artificial intelligence, cloud computing, digital fabrication, etc. These technologies help in developing machines that will be self-optimized and self-configured. It helps in enhancing asset performance, technology usage, material usage and other industrial processes that are involves in various industries. Numerous benefits are offered by these technologies such as low operational cost, improved productivity, enhanced customer satisfaction, improved customization, and increased efficiency. The Industry 4.0 holds a lot of potentials and is expected to register a substantial growth in the near future. There are several conferences on automation, robotics and communications, but they are not meet the Industry 4.0 challenges.

The series of annual ARCI Winter IFSA conferences have been launched to fill-in this gap and provide a forum for open discussion of state-of-the-art technologies related to control, automation, robotics and communication - three main components of Industry 4.0. It will be also to discuss how to adopt the current R&D results for Industry 4.0 and to customize products under the conditions of highly flexible (mass-) production.

The conference is organized by the International Frequency Sensor Association (IFSA) - one of the major professional, non-profit association serving for sensor industry and academy more than 20 years, in technical cooperation with media partners – journals: *MDPI Sensors* (Switzerland), *Soft Measurements and Computing* (Russia) and magazine *Manufacturing Technologies Insights* (USA). The conference program provides an opportunity for researchers interested in signal processing and artificial intelligence to discuss their latest results and exchange ideas on the new trends.

I hope that these proceedings will give readers an excellent overview of important and diversity topics discussed at the conference.

We thank all authors for submitting their latest works, thus contributing to the excellent technical contents of the Conference. Especially, we would like to thank the individuals and organizations that worked together diligently to make this Conference a success, and to the members of the International Program Committee for the thorough and careful review of the papers. It is important to point out that the great majority of the efforts in organizing the technical program of the Conference came from volunteers.

Prof., Dr. Sergey Y. Yurish
ARCI' 2021 Conference Chairman

(001)

Turbo Coded Single User Massive MIMO with Precoding

K. Vasudevan, Gyanesh Kumar Pathak, A. Phani Kumar Reddy
Department of Electrical Engineering, Indian Institute of Technology Kanpur-208016, India
{vasu, pathak, phani}@iitk.ac.in

Summary: Precoding is a method of compensating the channel at the transmitter. This work presents a novel method of data detection in turbo coded single user massive multiple input multiple output (MIMO) systems using precoding. We show via computer simulations that, when precoding is used, re-transmitting the data does not result in significant reduction in bit-error-rate (BER), thus increasing the spectral efficiency, compared to the case without precoding. Moreover, increasing the number of transmit and receive antennas results in improved BER.

Keywords: Precoding, Massive MIMO, Turbo codes, Flat fading, Spectral efficiency.

1. Introduction

Precoding at the transmitter is a technique that dates back to the era of voice band modems or wired communications [1-7]. The term "precoding" is quite generic and refers to one or more of the many different functionalities, as given below:

1. It compensates for the distortion introduced by the channel. Note that channel compensation at the receiver is referred to as equalization [8-14]. Here, channel compensation implies removal or minimization of intersymbol interference (ISI).
2. It performs error control coding, besides channel compensation.
3. It shapes the spectrum of the transmitted signal, and renders it suitable for propagation over the physical channel. Note that most channels do not propagate a DC signal and precoding is used to remove the DC component in the message signal. At this point, it is important to distinguish between a message signal and the transmitted signal.

In the context of wireless multiple input, multiple output (MIMO) systems, the main task of the precoder is to remove interchannel interference (ICI), either for single-user or multi-user case [15-21]. It should be observed that precoding requires knowledge of the channel state information (CSI) at the transmitter, which is usually fed back by the receiver to the transmitter. The receiver estimates CSI from a known training signal that is sent by the transmitter. CSI usually refers to the channel impulse response (CIR) or its statistics (mean and covariance), depending on the type of precoder used. Thus, precoding requires the channel to be time invariant or wide sense stationary (WSS) over at least one transmit and receive duration. Moreover, precoding can only be performed on systems employing time division duplex (TDD), which is a method of half duplex telecommunication. In other words, the channel needs to be reciprocal, that is, the CIR from the transmitter to receiver must be identical to that from receiver to transmitter.

In this work, we describe an elegant precoding method which reduces ICI in single user massive MIMO systems and compare it with the case without precoding [22, 23]. Rayleigh flat fading channel is assumed. If the channel is frequency selective, orthogonal frequency division multiplexing (OFDM) can be used [14, 23-32].

This work is organized as follows. Section 2 describes the signal model. In Section 3 precoding for single user massive MIMO is discussed. Section 4 presents the simulation results and conclude the work in Section 5.

2. Signal Model

Consider a precoded MIMO system with N_t transmit and N_r receive antennas, as shown in Fig. 1 [22]. The precoded received signal in the k^{th} ($0 \leq k \leq N_{rt} - 1$, k is an integer), re-transmission is given by

$$\widetilde{\mathbf{R}}_k = \widetilde{\mathbf{H}}_k \widetilde{\mathbf{H}}_k^H \mathbf{S}^p + \widetilde{\mathbf{W}}_k, \quad (1)$$

where $\widetilde{\mathbf{R}}_k \in \mathbb{C}^{N_r \times 1}$ is the received vector, $\widetilde{\mathbf{H}}_k \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix and $\widetilde{\mathbf{W}}_k \in \mathbb{C}^{N_r \times 1}$ is the additive white Gaussian noise (AWGN) vector. The transmitted symbol vector is $\mathbf{S}^p \in \mathbb{C}^{N_r \times 1}$, whose elements are drawn from an M-ary constellation. Boldface letters denote vectors or matrices. Complex quantities are denoted by a tilde. However tilde is not used for complex symbols $\mathbf{S}^p$. The elements of $\widetilde{\mathbf{H}}_k$ are statistically independent, zero mean, circularly symmetric complex Gaussian with variance per dimension equal to σ_H^2, as given by (2) of [22]. Similarly, the elements of $\widetilde{\mathbf{W}}_k$ are statistically independent, zero mean, circularly symmetric complex Gaussian with variance per dimension equal to σ_W^2, as given by (3) of [22].

In this work, the elements of $\mathbf{S}^p$ are turbo coded and mapped to a QPSK constellation with coordinates $\pm 1 \pm \mathrm{j}$, as depicted in Fig. 1. Moreover, here $\widetilde{\mathbf{H}}_k$ is an $N_r \times N_t$ matrix, whereas in [22] $\widetilde{\mathbf{H}}_k$ is an $N \times N$ matrix. We assume that $\widetilde{\mathbf{H}}_k$ and $\widetilde{\mathbf{W}}_k$ are independent across re-transmissions, hence (4) in [22] is valid with N replaced by N_r. We now proceed to analyze the signal model in (1).

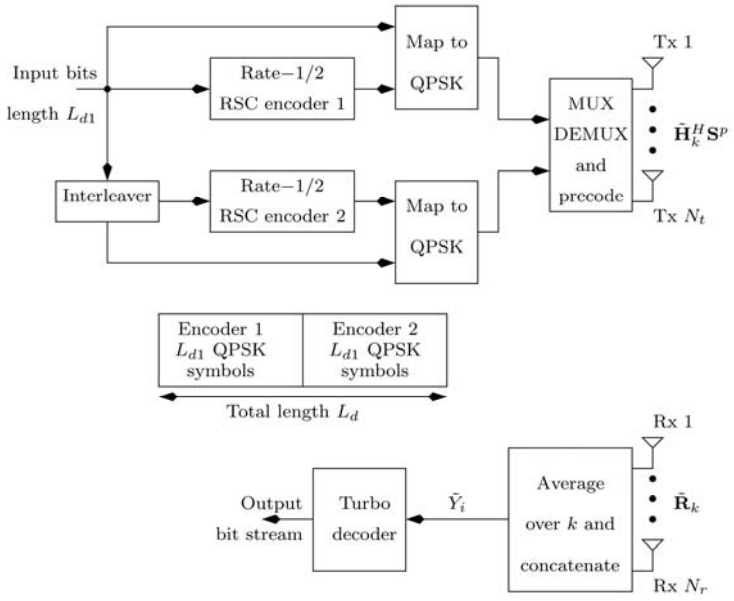


Fig. 1. System Model.

3. Precoding

The i^{th} element of $\widetilde{\mathbf{R}}_k$ in (1) is

$$\tilde{R}_{k,i} = \tilde{F}_{k,i,i} S_i + \tilde{I}_{k,i} + \tilde{W}_{k,i} \quad \text{for } 1 \le i \le N_r, \tag{2}$$

where

$$\tilde{F}_{k,i,i} = \sum_{j=1}^{N_t} \left|\widetilde{H}_{k,i,j}\right|^2, \tilde{I}_{k,i} = \sum_{\substack{j=1 \\ j\neq i}}^{N_r} \tilde{F}_{k,i,j} S_j, \quad \tilde{F}_{k,i,j} = \sum_{l=1}^{N_t} \widetilde{H}_{k,i,l} \widetilde{H}_{k,j,l}^* \quad \text{for } i \neq j \tag{3}$$

The desired signal in (2) is $F_{k,i,i}S_i$, the interference term is $\tilde{I}_{k,i}$ and the noise term is $\widetilde{W}_{k,i}$. Now

$$E\left[\tilde{F}_{k,i,i}^2\right] = E\left[\sum_{j=1}^{N_t} \left|\widetilde{H}_{k,i,j}\right|^2 \sum_{l=1}^{N_t} \left|\widetilde{H}_{k,i,l}\right|^2\right] = \\ = E\left[\sum_{j=1}^{N_t} \widetilde{H}_{k,i,j,I}^2 + \widetilde{H}_{k,i,j,Q}^2 \times \right. \\ \left. \times \sum_{l=1}^{N_t} \widetilde{H}_{k,i,l,I}^2 + \widetilde{H}_{k,i,l,Q}^2\right] = \\ = 4\sigma_H^4 N_t (N_t + 1), \tag{4}$$

where the subscript "I" denotes the in-phase part and the subscript "Q" denotes the quadrature part of a complex quantity and the following relation has been used [33, 34]

$$E[X^4] = 3\sigma_X^4, \tag{5}$$

where X is a zero-mean, real-valued Gaussian random variable with variance σ_X^2. Moreover from (3) and (2) in [22]

$$E\left[\tilde{F}_{k,i,i}\right] = 2\sigma_H^2 N_t \tag{6}$$

We also have

$$E\left[\left|\tilde{I}_{k,i}\right|^2\right] = E\left[\sum_{\substack{j=1 \\ j\neq i}}^{N_r} \tilde{F}_{k,i,j} S_j \times \right. \\ \left. \times \sum_{\substack{l=1 \\ l\neq i}}^{N_r} \tilde{F}_{k,i,l}^* S_l^*\right] = \\ = \sum_{\substack{j=1 \\ j\neq i}}^{N_r} \sum_{\substack{l=1 \\ l\neq i}}^{N_r} P_{\text{av}} E\left[\tilde{F}_{k,i,j} \tilde{F}_{k,i,l}^*\right] \delta_K(j-l) = \\ = P_{\text{av}} \sum_{\substack{j=1 \\ j\neq i}}^{N_r} E\left[\left|\tilde{F}_{k,i,j}\right|^2\right], \tag{7}$$

where $\delta_K(\cdot)$ is the Kronecker delta function [14, 22], we have assumed independence between $\tilde{F}_{k,i,j}$ and S_j and [22]

$$E\left[S_j S_l^*\right] = P_{\text{av}} \delta_K(j-l) = 2\delta_K(j-l) \tag{8}$$

Now

$$E\left[\left|\tilde{F}_{k,i,j}\right|^2\right] = E\left[\sum_{l=1}^{N_t} \widetilde{H}_{k,i,l} \widetilde{H}_{k,j,l}^* \times \right. \\ \left. \times \sum_{m=1}^{N_t} \widetilde{H}_{k,i,m}^* \widetilde{H}_{k,j,m}\right] = \\ = \sum_{l=1}^{N_t} \sum_{m=1}^{N_t} E\left[\widetilde{H}_{k,i,l} \widetilde{H}_{k,i,m}^*\right] \times \\ \times E\left[\widetilde{H}_{k,j,m} \widetilde{H}_{k,j,l}^*\right] = \\ = \sum_{l=1}^{N_t} \sum_{m=1}^{N_t} 4\sigma_H^4 \delta_K(l-m) = 4\sigma_H^4 N_t \tag{9}$$

Substituting (9) in (7) and using (8) we get

$$E\left[\left|\tilde{I}_{k,i}\right|^2\right] = 8\sigma_H^4 N_t (N_r - 1) \tag{10}$$

Due to independence between $\tilde{I}_{k,i}$ and $\widetilde{W}_{k,i}$ in (2) we have from (10) and (3) of [22]

$$E\left[\left|\tilde{I}_{k,i} + \widetilde{W}_{k,i}\right|^2\right] = E\left[\left|\tilde{I}_{k,i}\right|^2\right] + \\ + E\left[\left|\widetilde{W}_{k,i}\right|^2\right] = 8\sigma_H^4 N_t (N_r - 1) + 2\sigma_W^2 = \\ = \sigma_{U'}^2 \text{ (say)} \tag{11}$$

Now, each element of $\mathbf{S}^p$ in (1) carries $1/(2N_{rt})$ bits of information [22]. Therefore, each element of $\widetilde{\mathbf{R}}_k$ also carries $1/(2N_{rt})$ bits of information. Hence, the average signal to interference plus noise ratio per bit of $\tilde{R}_{k,i}$ in (2) is defined as, using (4), (8) and (11)

$$\begin{aligned}\mathrm{SINR}_{\mathrm{av},b} &= \frac{E\left[\left|\tilde{F}_{k,i,i}S_i\right|^2\right]\times 2N_{rt}}{E\left[\left|I_{k,i}+\widetilde{W}_{k,i}\right|^2\right]} = \\ &= \frac{8\sigma_H^4 N_t(N_t+1)\times 2N_{rt}}{8\sigma_H^4 N_t(N_r-1)+2\sigma_W^2}\end{aligned} \tag{12}$$

When $\sigma_W^2 = 0$ in (12), we get the upper bound on $\mathrm{SINR}_{\mathrm{av},b}$ as given below

$$\begin{aligned}\mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}} &= \frac{8\sigma_H^4 N_t(N_t+1)\times 2N_{rt}}{8\sigma_H^4 N_t(N_r-1)} = \\ &= \frac{2N_{rt}(N_t+1)}{N_r-1}\end{aligned} \tag{13}$$

The information contained in $\mathbf{S}^p$ in (1) is $N_r/(2N_{rt})$ bits. Hence the spectral efficiency of the precoded system is

$$\eta^p = \frac{N_r}{2N_{rt}} \text{ bits per transmission} \tag{14}$$

Note that both (13) and (14) need to be as large as possible to minimize the BER and maximize the spectral efficiency. Let

$$N_{\mathrm{tot}} = N_t + N_r \tag{15}$$

Define

$$\begin{aligned}f(N_t) &= \mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}} + \eta^p = \\ &= \frac{2N_{rt}(N_t+1)}{N_r-1} + \frac{N_r}{2N_{rt}} = \frac{2N_{rt}(N_t+1)}{N_{\mathrm{tot}}-N_t-1} + \frac{N_{\mathrm{tot}}-N_t}{2N_{rt}},\end{aligned} \tag{16}$$

where we have used (15). We need to find N_t such that $f(N_t)$ is maximized. The plot of $\mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}}$ (red curve), η^p (blue curve) and $f(N_t)$ (green curve), as a function of N_t, keeping N_{tot} fixed, is depicted in Fig. 2 and **3**. Note that $\mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}}$ increases monotonically and η^p decreases monotonically, with increasing N_t. We also find that $f(N_t)$ has a minimum (not maximum) at

$$N_t = N_{\mathrm{tot}} - 2N_{rt}\sqrt{N_{\mathrm{tot}}} - 1, \tag{17}$$

which is obtained by differentiating $f(N_t)$ in (16) with respect to N_t and setting the result to zero. Therefore, the only possible solution is to avoid the minimum. Clearly we require $\mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}} > \ln(2)$, since it is the minimum average SNR per bit required for error-free transmission over any type of channel [22]. We also require $\eta^p > \eta^p_{\min}$, where $\eta^p_{\min}$ is chosen by the system designer. Thus, we arrive at a range of the number of transmit antennas ($N_{t,\min} \le N_t \le N_{t,\max}$) that can be used, as shown in Fig. 2 and 3. Note that in Fig. 3(b) the minimum of $f(N_t)$ cannot be avoided, since $\eta^p_{\min}$ would be too small.

Next, similar to (20) in [22], consider

$$\begin{aligned}\tilde{Y} &= \frac{1}{N_{rt}}\sum_{k=0}^{N_{rt}-1}\tilde{R}_{k,i} = \\ &= \frac{1}{N_{rt}}\sum_{k=0}^{N_{rt}-1}\left(\tilde{F}_{k,i,i}S_i + \tilde{I}_{k,i} + \widetilde{W}_{k,i}\right) = \\ &= F_i S_i + \widetilde{U}_i \text{ for } 1 \le i \le N_r,\end{aligned} \tag{18}$$

where $\tilde{R}_{k,i}$ is given by (2), F_i is real-valued and

$$\begin{aligned}F_i &= \frac{1}{N_{rt}}\sum_{k=0}^{N_{rt}-1}\tilde{F}_{k,i,i}, \\ \widetilde{U}_i &= \frac{1}{N_{rt}}\sum_{k=0}^{N_{rt}-1}\left(\tilde{I}_{k,i} + \widetilde{W}_{k,i}\right) = \\ &= \frac{1}{N_{rt}}\sum_{k=0}^{N_{rt}-1}\widetilde{U}_{k,i}' \text{ (say)}\end{aligned} \tag{19}$$

Since $\tilde{F}_{k,i,i}$ and $\widetilde{U}_{k,i}'$ are statistically independent over re-transmissions (k), we have

$$\begin{aligned}E[F_i^2] &= \frac{1}{N_{rt}^2}E\left[\sum_{k=0}^{N_{rt}-1}\tilde{F}_{k,i,i}\sum_{n=0}^{N_{rt}-1}\tilde{F}_{n,i,i}\right] = \\ &= \frac{4\sigma_H^4 N_t[N_t+1+N_t(N_{rt}-1)]}{N_{rt}} = \frac{4\sigma_H^4 N_t(N_tN_{rt}+1)}{N_{rt}}, \\ E\left[\left|\widetilde{U}_i\right|^2\right] &= \frac{\sigma_{U'}^2}{N_{rt}} = \frac{8\sigma_H^4 N_t(N_r-1)+2\sigma_W^2}{N_{rt}},\end{aligned} \tag{20}$$

where we have used (4), (6), (11) and the fact that

$$E\left[\widetilde{U}_{k,i}'\right] = 0, \tag{21}$$

where $\widetilde{U}_{k,i}'$ is defined in (19). Next, we compute the average SINR per bit for $\tilde{Y}_i$ in (18). Note that since $\tilde{Y}_i$ is a "combination" of N_{rt} re-transmissions, its information content is $N_{rt}/(2N_{rt}) = 1/2$ bit (recall that the information content of $\tilde{R}_{k,i}$ in (18) is $1/(2N_{rt})$ bits). Therefore

$$\mathrm{SINR}_{\mathrm{av},b,C} = \frac{E\left[|F_iS_i|^2\right]\times 2}{E\left[|\widetilde{U}_i|^2\right]} = \frac{8\sigma_H^4 N_t(N_tN_{rt}+1)\times 2}{8\sigma_H^4 N_t(N_r-1)+2\sigma_W^2}, \tag{22}$$

where the subscript "C" denotes "after combining" and we have used (8) and (20). Note that we prefer to use the word "combining" rather than averaging, since it is more appropriate in terms of the "information content" in $\tilde{Y}_i$. Once again with $\sigma_W^2 = 0$ and $N_tN_{rt} \gg 1$ we get the approximate upper bound on $\mathrm{SINR}_{\mathrm{av},b,C}$ as

$$\begin{aligned}\mathrm{SINR}_{\mathrm{av},b,C,\mathrm{UB}} &= \frac{8\sigma_H^4 N_t(N_tN_{rt}+1)\times 2}{8\sigma_H^4 N_t(N_r-1)} \approx \frac{2N_{rt}N_t}{N_r-1} \approx \\ &\approx \mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}},\end{aligned} \tag{23}$$

when $N_t \gg 1$. Thus, the upper bound on the average SINR per bit before and after combining are nearly identical. Observe that re-transmitting the data increases the upper bound on the average SINR per bit, it does not improve the BER performance, which is seen in the next section. After concatenation, the signal $\tilde{Y}_i$ in (18) for $0 \le i \le L_d - 1$ is sent to the turbo decoder. The details of turbo decoding will not be discussed here.

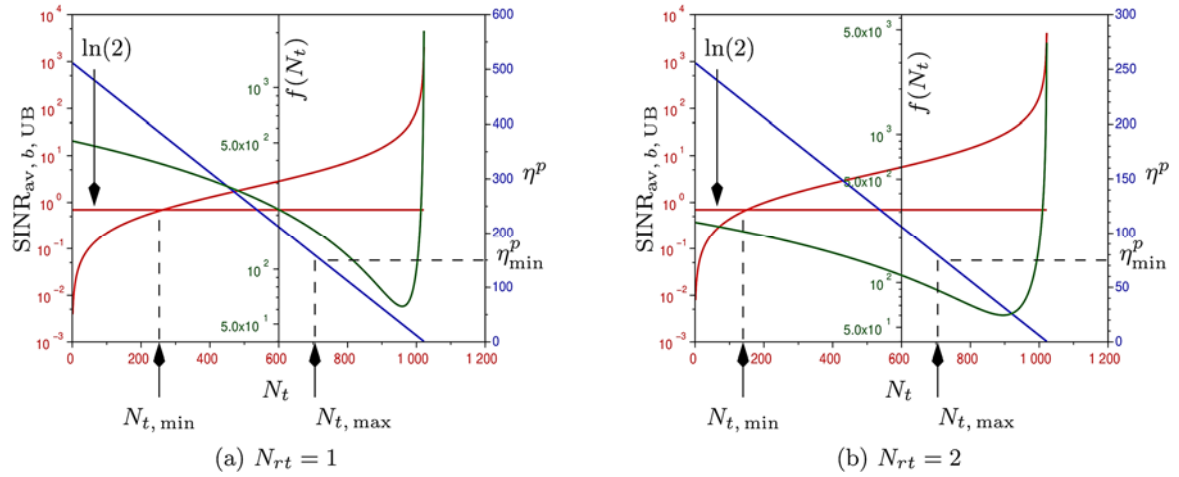


Fig. 2. $\mathrm{SINR}_{\mathrm{av},b,\mathrm{UB}}$ and η^p as a function of N_t for $N_{\mathrm{tot}} = 1024$.

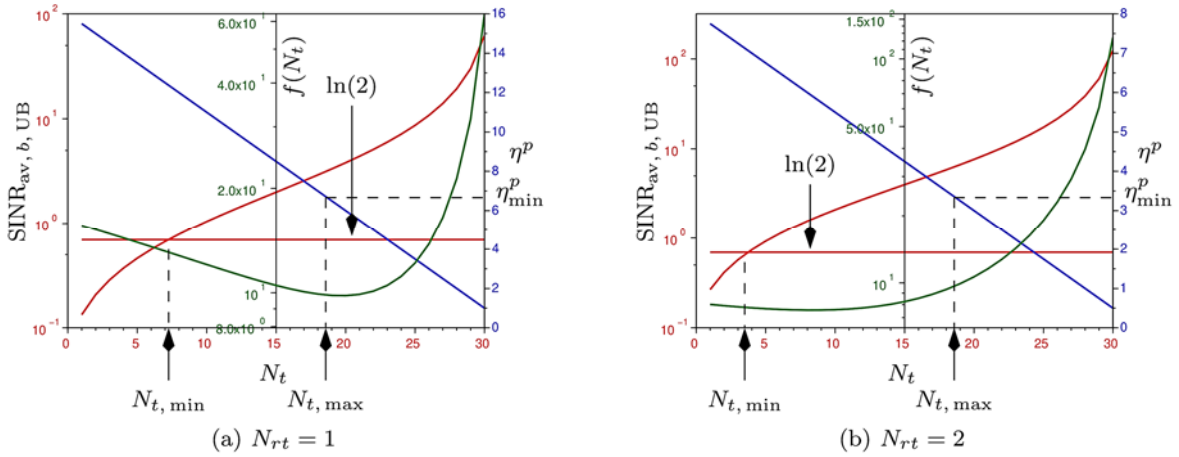


Fig. 3. SINRav, b, UB and η^p as a function of N_t for $N_{\mathrm{tot}} = 32$.

4. Simulation Results

In this section, we discuss the results from computer simulations. The length of the data bits per "frame" (L_{d1}) is taken to be the smallest integer greater than 1000, which is an integer multiple of N_r. Note that (see Fig. 1)

$$L_d = 2L_{d1} \tag{24}$$

The simulations were carried out over 10^4 frames. The turbo encoder is given by (38) of [22].

Fig. 4(a) gives the bit-error-rate (BER) results for a 1×1 single input single output (SISO) system ($N_{\mathrm{tot}} = 2$). We get a BER of 2×10^{-2} at an average SNR per bit of 3.5 dB, with $N_{rt} = 4$. The corresponding spectral efficiency is $\eta^p = 1/8$ bits per transmission. Th BER also does not vary significantly with the number of re-transmissions (N_{rt}).

Fig. 4 (b) gives the results for $N_{\mathrm{tot}} = 32$ and different combinations of transmit (N_t) and receive (N_r) antennas. We find that the BER is quite insensitive to variations in N_t, N_r and N_{rt}. Moreover, the BER at an SNR per bit of 3.5 dB is about 2×10^{-6}, which is a significant improvement over the SISO system. Of all the curves, $N_t = 25$, $N_{rt} = 2$ gives the lowest spectral efficiency of $\eta^p = 1.75$ bits/sec/Hz and highest $\mathrm{SNR}_{\mathrm{av},b,\mathrm{UB}} = 12.39$ dB. Of all the curves, $N_t = 12$, $N_{rt} = 1$ gives the highest spectral efficiency $\eta^p = 10$ bits/sec/Hz and lowest $\mathrm{SNR}_{\mathrm{av},b,\mathrm{UB}} = 1.36$ dB.

Fig. 4(c) gives the results for $N_{\mathrm{tot}} = 1024$ for various combinations of N_t, N_r and N_{rt}. The BER is similar to that of $N_{\mathrm{tot}} = 32$. Of all the curves, $N_t = 400$, $N_{rt} = 1$ gives the highest spectral efficiency of $\eta^p = 312$ bits/sec/Hz and lowest $\mathrm{SNR}_{\mathrm{av},b,\mathrm{UB}} = 1.09$ dB. Of all the curves, $N_t = 1023$, $N_{rt} = 2$ gives the lowest spectral efficiency of $\eta^p = 0.25$ and highest $\mathrm{SNR}_{\mathrm{av},b,\mathrm{UB}} \to \infty$.

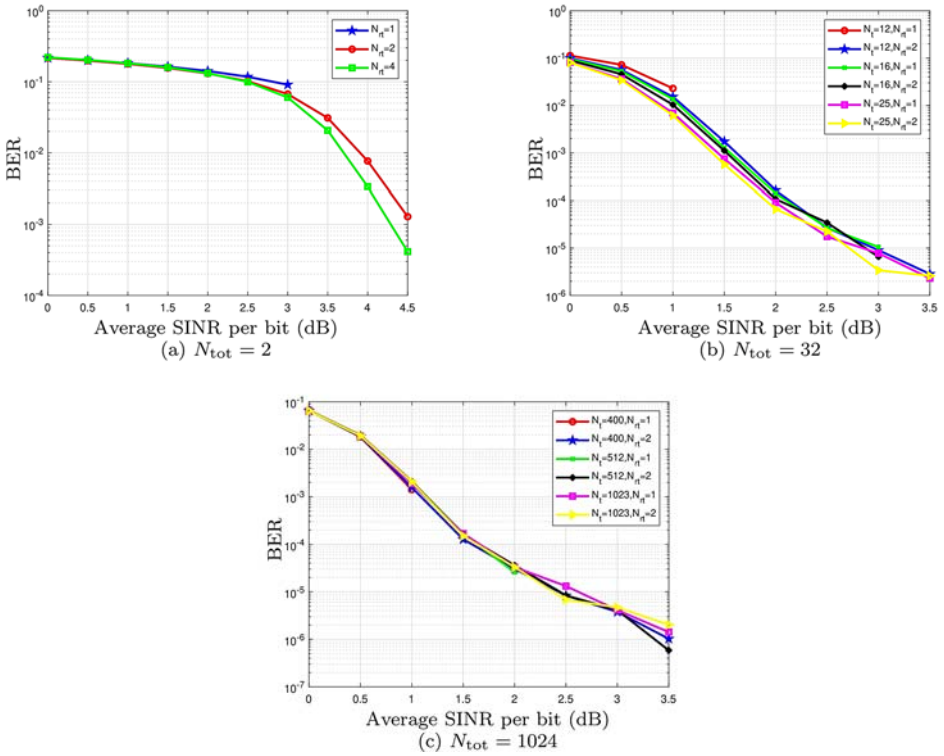


Fig. 4. Simulation Results.

5. Conclusions

This work presents a method for data detection in turbo-coded and precoded massive MIMO. An ideal receiver is assumed. Future work could be to simulate a realistic precoded system with carrier and timing synchronization and channel estimation.

References

[1]. R. W. Chang, Precoding for multiple-speed data transmission, *The Bell System Technical Journal*, Vol. 46, Issue 7, Sep. 1967, pp. 1633-1649.

[2]. S. Kasturia, J. M. Cioffi, Precoding for blocking signaling and shaped signal sets, in *Proceedings of the IEEE International Conference on Communications, World Prosperity Through Communications*, Vol. 2, June 1989, pp. 1086-1090.

[3]. G. J. Pottie, M. V. Eyuboglu, Combined coding and precoding for PAM and QAM HDSL systems, *IEEE Journal on Selected Areas in Communications*, Vol. 9, Issue 6, Aug 1991, pp. 861-870.

[4]. A. K. Aman, R. L. Cupo, N. A. Zervos, Combined trellis coding and DFE through Tomlinson precoding, *IEEE Journal on Selected Areas in Communications*, Vol. 9, Issue 6, Aug 1991, pp. 876-884.

[5]. G. D. Forney, M. V. Eyuboglu, Combined equalization and coding using precoding, *IEEE Communications Magazine*, Vol. 29, Issue 12, Dec 1991, pp. 25-34.

[6]. M. V. Eyuboglu, G. D. Forney, Trellis precoding: combined coding, precoding and shaping for intersymbol interference channels, *IEEE Transactions on Information Theory*, Vol. 38, Issue 2, 1992, pp. 301-314.

[7]. R. Laroia, S. A. Tretter, N. Farvardin, A simple and effective precoding scheme for noise whitening on intersymbol interference channels, *IEEE Transactions on Communications*, Vol. 41, Issue 10, 1993, pp. 1460-1463.

[8]. H. W. Bode, Variable equalizers, *The Bell System Technical Journal*, Vol. 17, Issue 2, 1938, pp. 229-244.

[9]. S. U. H. Qureshi, Adaptive equalization, *Proceedings of the IEEE*, Vol. 73, Issue 9, 1985, pp. 1349-1387.

[10]. K. Vasudevan, Detection of signals in correlated interference using a predictive VA, in *Proceedings of the 8th International Conference on Communication Systems (ICCS'02)*, Vol. 1, 2002, pp. 529-533.

[11]. K. Vasudevan, Detection of signals in correlated interference using a predictive VA, *Signal Processing*, Vol. 84, Issue 12, 2004, pp. 2271-2286.

[12]. R. Koetter, A. C. Singer, M. Tüchler, Turbo Equalization, *IEEE Sig. Proc. Mag.*, Vol. 21, Issue 1, Jan. 2004, pp. 67-80.

[13]. K. Vasudevan, Turbo equalization of serially concatenated turbo codes using a predictive DFE-based receiver, *Signal, Image and Video Processing*, Vol. 1, Issue 3, Aug. 2007, pp. 239-252.

[14]. K. Vasudevan, Digital Communications and Signal Processing, 2nd Edition, *Universities Press (India)*, Hyderabad, 2010.

[15]. N. Fatema, G. Hua, Y. Xiang, D. Peng, I. Natgunanathan, Massive MIMO linear precoding: A survey, *IEEE Systems Journal*, Vol. 12, Issue 4, Dec. 2018, pp. 3920-3931.

[16]. S. Kim, Diversity order of precoding-aided spatial modulation using receive antenna selection, *Electronics Letters*, Vol. 56, Issue 5, 2020, pp. 260-262.

[17]. S. Kim, Transmit antenna selection for precoding-aided spatial modulation, *IEEE Access*, Vol. 8, 2020, pp. 40723-40731.

[18]. A. Haqiqatnejad, F. Kayhan, B. Ottersten, Robust SINR-constrained symbol-level multiuser precoding with imperfect channel knowledge, *IEEE Transactions on Signal Processing*, Vol. 68, pp. 1837-1852, 2020.

[19]. Q. Deng, X. Liang, X. Wang, M. Huang, C. Dong, Y. Zhang, Fast converging iterative precoding for massive MIMO systems: An accelerated weighted Neumann series-steepest descent approach, *IEEE Access*, Vol. 8, 2020, pp. 50244-50255.

[20]. Z. Li, C. Zhang, I. Lu, X. Jia, Hybrid precoding using out-of-band spatial information for multi-user multi RF chain millimeter wave systems, *IEEE Access*, Vol. 8, 2020, pp. 50872-50883.

[21]. N. L. Johannsen, N. Peitzmeier, P. A. Hoeher, D. Manteuffel, On the feasibility of multi-mode antennas in UWB and IoT applications below 10 GHz, *IEEE Communications Magazine*, Vol. 58, Issue 3, 2020, pp. 69-75.

[22]. K. Vasudevan, K. Madhu, S. Singh, data detection in single user massive MIMO using re-transmissions, *The Open Signal Processing Journal*, Vol. 6, Mar. 2019, pp. 15-26.

[23]. K. Vasudevan, S. Singh, A. P. K. Reddy, Coherent receiver for turbo coded single-user massive MIMO OFDM with retransmissions, Chapter 4, in Multiplexing (S. Mohammady, Ed.), *IntechOpen*, London, 2019, pp. 1-21.

[24]. K. Vasudevan, Coherent detection of turbo coded OFDM signals transmitted through frequency selective Rayleigh fading channels, in *Proceedings of the IEEE International Conference on Signal Processing, Computing and Control (ISPCC'13)*, Sept. 2013, pp. 1-6.

[25]. K. Vasudevan, Coherent detection of turbo-coded OFDM signals transmitted through frequency selective Rayleigh fading channels with receiver diversity and increased throughput, *Wireless Personal Communications*, Vol. 82, Issue 3, 2015, pp. 1623-1642.

[26]. K. Vasudevan, Coherent turbo coded MIMO OFDM, in *Proceedings of the 12th International Conference on Wireless and Mobile Communications (ICWMC'16)*, Nov. 2016, pp. 91-99.

[27]. K. Vasudevan, Near capacity signaling over fading channels using coherent turbo coded OFDM and massive MIMO, *International Journal On Advances in Telecommunications*, Vol. 10, Issues 1-2, 2017, pp. 22-37.

[28]. K. Vasudevan, SCILAB code for coherent detection of turbo coded OFDM signals transmitted through frequency selective Rayleigh fading channels, https://www.codeocean.com/

[29]. K. Vasudevan, K. Madhu, S. Singh, SCILAB code for data detection in single user massive MIMO using retransmissions, https://www.codeocean.com/

[30]. K. Vasudevan, S. Singh, A. P. K. Reddy, SCILAB code for coherent receiver for turbo coded single user massive MIMO-OFDM with retransmissions, https://www.codeocean.com/

[31]. K. Vasudevan, A. Phani Kumar Reddy, G. K. Pathak, S. Singh, On the probability of erasure for MIMO OFDM, *Semiconductor Science and Information Devices*, Vol. 2, Issue 1, Apr. 2020, pp. 1-5.

[32]. K. Vasudevan, A. P. K. Reddy, G. K. Pathak, S. Singh, SCILAB code for the probability of erasure for MIMO-OFDM, https://www.codeocean.com/

[33]. A. Papoulis, Probability, Random Variables and Stochastic Processes, 3rd Edition, *McGraw-Hill*, 1991.

[34]. K. Vasudevan, Analog Communications: Problems & Solutions, *Ane Books*, 2018.

(002)

Wearable Sensor Technology for Individual Grip Force Profiling

Rongrong Liu [1], Florent Nageotte [1], Philippe Zanne [1], Michel de Mathelin [1] and Birgitta Dresp-Langley [2]

[1] ICube Lab, Robotics Department, University of Strasbourg, UMR 7357, CNRS, 67000 Strasbourg, France
[2] ICube Lab, UMR 7357, Centre National de la Recherche Scientifique CNRS, 67000 Strasbourg, France
E-mails: rong.rong.liu@unistra.fr, Nageotte@unistra.fr, zanne.philippe@unistra.fr, demathelin@unistra.fr, birgitta.dresp@unistra.fr

Summary: Wearable biosensor systems with transmitting capabilities represent innovative technology developed to monitor exercise and other task activities. This technology enables real-time, convenient, and continuous monitoring of a user's behavioral signals, relative to body motion, body temperature and a variety of biological or biochemical markers, like individual grip force, which is studied in this paper. To achieve this goal, a four-step pick-and-drop image-guided robot-assisted precision task has been designed using a wearable wireless sensor glove system. The spatio-temporal grip force profiling is analyzed on the basis of thousands of individual sensor data collected from the twelve locations on the dominant and non-dominant hands of each of the three users in ten successive task sessions. Statistical comparison have shown specific differences between the grip force profiles of individual users as a function of task skill level and expertise.

Keywords: Wearable biosensor technology, Individual grip force, Image-guided task, Robot-assisted task, Wearable wireless sensor glove system, Spatio-temporal profiling.

1. Introduction

Wearable sensors, as the name implies, are integrated into wearable objects or directly with the body in order to monitor and transmit a user's behavioral signals in real time. In this paper, the spatio-temporal grip force profiling will be studied based on the data collected from a wearable wireless sensor glove system developed in the lab [1, 2].

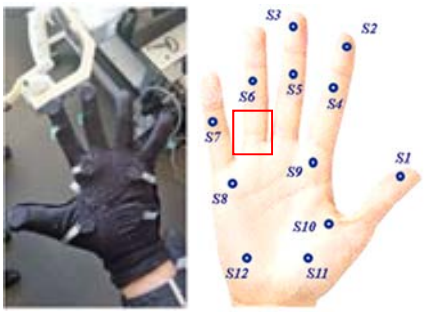
2. Materials and Methods

A specific wearable sensor system in terms of a glove for each hand with inbuilt Force Sensitive Resistors (FSR) has been developed. The hardware and software configurations will be briefly described here below. For further detailed information, one may go to https://www.mdpi.com/1424-8220/19/20/4575/htm.

The wireless sensor glove hardware-software system, is designed for bi-manual intervention, and task simulations may solicit either the dominant or the non-dominant hand, or both hands at the same time. For each hand, twelve anatomically relevant FSR are employed to measure the grip force applied on certain locations on the fingers and in the palm. These FSR have been sewn into a soft glove (Fig. 1(a)) and their locations are shown in Fig. 1(b).

The software of the glove system includes two parts: one running on the gloves, and the other running on the computer algorithm for data collection. During the experiment, each of the two gloves is sending data to the computer separately every 20 milliseconds (50 Hz), merged with the time stamps and sensor identification. This data package is sent to the computer via Bluetooth, which will be then decoded and saved by the computer software.

A four-step pick-and-drop image-guided robot-assisted precision task, as described in Table 1, has been designed for this individual grip force study. Grip force data are analyzed here for one left-handed highly proficient expert, and one right-handed complete novice.



(a) (b)

Fig. 1. Signals relative to grip force are sampled from 12 anatomically relevant FSR locations on the fingers and in the palm of both hands.

Table 1. Four-step pick-and-drop task.

Step	Description
1	Activate and move tool towards object location
2	Open and close grippers to grasp and lift object
3	Move tool with object to target location
4	Open grippers to drop object in box

3. Results

Several thousands of grip force data have been collected from the twelve sensor locations in ten

successive sessions for repeated execution of the pick-and-drop robotic task. The corresponding task times for the dominant and non-dominant hands of each of the three users are illustrated in Table 2.

Table 2. Task time (sec).

User	Dominant	Non-dominant
Expert	8.88	10.19
Trained	11.90	13.53
Novice	15.42	12.99

As the middle phalanx of the small finger allows for precision grip force control [3, 4], and is critically important in surgical and other precision tasks, we focus on the corresponding sensor S7 on the dominant hand of expert and novice as the most representative.

Individual spatio-temporal grip force profiles have been plotted in Fig. 2 for the first and the last individual sessions, in terms of average peak amplitudes (mV) for fixed successive temporal windows of 2000 milliseconds (100 signals per time window, session, and user).

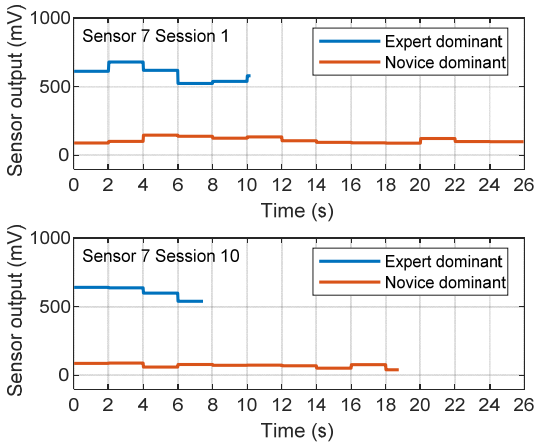


Fig. 2. Individual spatio-temporal grip force profiles showing average peak amplitudes (mV) from sensor S7 for fixed successive temporal windows of 2000 milliseconds for the first and last of ten sessions of two users.

A 2-Way ANOVA on the raw grip force data has been conducted for statistical comparison. The expertise-specific difference between the two user profiles is characterized by the novice deploying largely insufficient grip forces, from the first session with m = 98 mV /sem = 1.2 to the last with m = 78 mV/sem = 1.6, while the expert produces sufficient grip force for fine movement control from the first session with m = 594 mV/sem = 1.8 to the last with m = 609 mV/sem = 2.2. The interaction between the 'expertise' and 'session' factors for sensor S7 is highly significant with $F(1,2880) = 188.53$; $p < .001$.

4. Discussion

The spatio-temporal profiling and statistical comparison have shown specific differences between the grip force profiles of individual users as a function of task skill level and expertise in using the robotic system. Experts and non-experts employ different grip-force strategies, reflected by differences in amount of grip force deployed by the middle phalanx of the small finger, with the novice dominant hand deploying insufficient grip forces, and no major evolution between the first and the last task sessions.

In terms of task time, at the beginning, the novice takes more than twice as long performing the precision task by comparison with the expert, but at the end scores a 30 % time gain indicating a considerable temporal training effect.

5. Conclusions

Grip force analysis on wearable sensors signals is a powerful means of tracking the evolution of individual force profile. The analyses shown in this paper here can deliver insight to monitor manual/bimanual precision tasks, control performance quality, or prevent risks in robot-assisted surgery systems, where excessive grip forces can cause tissue damage [5].

Acknowledgements

Material support by CNRS is gratefully acknowledged by the authors.

References

[1]. M. de Mathelin, F. Nageotte, P. Zanne, B. Dresp-Langley, Sensors for expert grip force profiling: towards benchmarking manual control of a robotic device for surgical tool movements, *Sensors (Basel)*, Vol. 19, Issue 20, 2019, 4575.

[2]. A. U. Batmaz, A. M. Falek, L. Zorn, F. Nageotte, P. Zanne, M. de Mathelin, B. Dresp-Langley, Novice and expert behavior while using a robot controlled surgery system, in *Proceedings of the 13th IASTED International Conference on Biomedical Engineering (BioMed'17)*, Innsbruck, Austria, 20-21 February 2017, pp. 94-99.

[3]. M. L. Latash, V. M. Zatsiorsky, Multi-finger prehension: Control of a redundant mechanical system, *Adv. Exp. Med. Biol.*, Vol. 629, 2009, pp. 597-618.

[4]. H. Kinoshita, S. Kawai, K. Ikuta, Contributions and co-ordination of individual fingers in multiple finger prehension, *Ergonomics*, Vol. 38, Issue 6, 1995, pp. 1212-1230.

[5]. A. Abiri, J. Pensa, A. Tao, J. Ma, Y. Y. Juo, S. J. Askari, J. Bisley, J. Rosen, E. P. Dutson, W. S. Grundfest, Multi-modal haptic feedback for grip force reduction in robotic surgery, *Scientific Reports*, Vol. 9, Issue 1, 2019, 5016.

(003)

Robot-robot Cooperation for Efficient Drilling of Soft Materials

J. Gotlih, T. Karner, M. Ficko and M. Brezočnik
[1] University of Maribor, Laboratory of Intelligent Manufacturing Systems, Smetanova 17,
2000 Maribor, Slovenia
Tel.: + 386 2 220 7605
E-mail: janez.gotlih@um.si

Summary: In some machining applications industrial robots are shown as a valid alternative to specialized machine tools. For example, in case of machining large work pieces, soft materials, or when high tolerance levels are not required but high system flexibility is, and when overall cost is a concern, a robotic machining setup can be considered. In traditional machining, mounting a work piece to a machine is the least automated process, while in case of robotic machining a serving robot can easily be included. To increase the flexibility of robotic machining a robot-robot cooperation approach is proposed, where one robot is used for mounting and manipulating the work piece, and another is used for machining the work piece to the desired shape. For optimal execution of the process a genetic algorithm optimization is set up to search for the layout of both robots and a drilling trajectory, with the goal to maximize manipulability of the system during drilling. Optimization result are verified in a simulated environment. It is shown that, using the genetic algorithm, an optimal dual robot cell can be designed, regarding the task executed by the robots.

Keywords: Robotic cooperation, Robotic machining, Manipulability, Optimization, Genetic algorithm.

1. Introduction

Solutions where robots are in direct contact with each other are rare, although a flexible, reconfigurable, and fast production system suggests direct robotic cooperation. In automotive and assembly industry, where flexibility is subordinate to batch size and production speed to task sequencing, indirect robotic cooperation is common [1], but for smaller production runs, these system traits become more important. If jobs frequently change and machine quantity and utilization is a concern, an improvement in system flexibility and reconfigurability could be achieved by direct robotic cooperation.

The manufacturing technology where mentioned system traits are highly desirable is machining. Until today, robots were already successfully introduced to various machining operations, increasing competition in the current machining environment. The main advantages of using robots for machining are [2]:

- Price-competitive with traditional machine tools;
- Broad reach such as the deburring or polishing of Large parts for aircraft or wind turbines;
- Very high levels of precision afforded by machine tools are not needed in some sectors and applications,

while the main disadvantages are:

- Low levels of precision because of low stiffness;
- Speed in machining parts with long toolpaths.

A fast, flexible robotic machining system suggests robots in direct cooperation, consisting of a service robot for work piece mounting and manipulation and another robot for machining. Motion control by the use of linked motion allows one robot to link or unlink to a reference frame on another robot, while that robot is moving without stopping motion [3]. Such an approach was already efficiently implemented as a master-slave setup [4].

To increase the systems performance, functional redundancy can be exploited. For robotic systems kinematic performance measures like manipulability are employed [5]. Manipulability is used in:

- Robot design;
- Trajectory planning to avoid singularities;
- Optimization of robotic machining [6].

Optimization of robotic systems is computationally demanding, but several nondeterministic methods like the genetic algorithm have proven suitable [7]. The main advantage of the GA is its ability to cope with nonlinear problems, common in robotics. Up to now, GA was successfully applied for robot topology optimization, robot calibration, work piece placement and tool trajectory planning. Recently, hybrid GA was found to be the most effective algorithm when applied to multi-robot cellular manufacturing systems [8], while swarm algorithms were considered for multi-robot welding path positioning [9].

For optimal cooperation of robots during robotic drilling relative positions between the cooperating robots and the drilling path need to be determined. For simultaneous path placement and trajectory planning a high degree of functional redundancy must be overcome, which was already efficiently solved by a kinematic optimization for wrapping a work piece [10]. For optimization of robotic drilling, also the robot's stiffness should be considered as it was found to have an important effect on hole quality [11, 12].

This article studies the feasibility of direct robotic cooperation for robotic drilling. A dual robot cell is set up, whereby the first robot is fixed in the global reference frame. The position of the second robot relative to the first robot's base frame and the drilling

trajectory position and orientation are the optimization variables. The system's total manipulability for executing the drilling task is defined as the objective function. Restrictions of the common workspace and simplifications of the collision body geometries were applied to increase the algorithm's performance and avoid undesirable solutions with equivalent fitness. A single-objective GA algorithm is applied to solve the problem. The final solution of the theoretical example is verified by an industrial software for offline programing of robots.

2. Materials and Methods

2.1. Robot Type and Technology

In this theoretical study, two ABB IRB 1200 5/0.9 robots (Fig. 1) were used to mount, manipulate, and drill holes into a soft polyurethane work piece. The finished part was a block by length $120\ mm$, width $120\ mm$, thickness $20\ mm$ and two through holes with diameter $\emptyset 4\ mm$. The service robot was used to pick up and mount the workpiece at the input buffer and position it into the drilling robot's workspace, where the drilling robot drilled the two holes into the workpiece by following a prescribed trajectory.

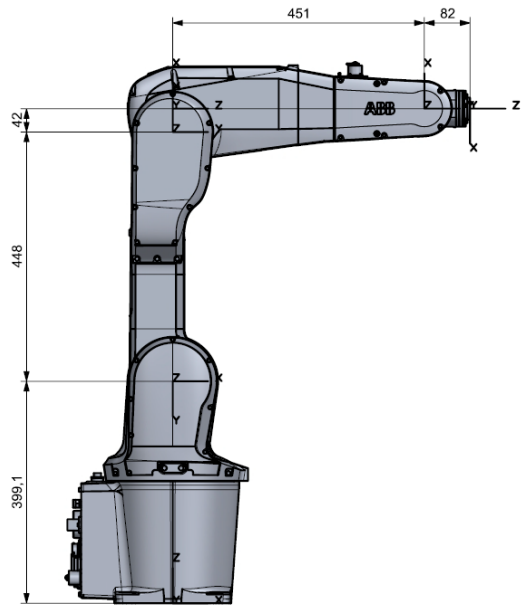


Fig. 1. ABB IRB 1200 5/0.9 robot with DH frames.

2.2. ABB IRB 1200 5/0.9 Robot Kinematic Model

The kinematic model of the ABB IRB 1200 5/0.9 robot was constructed by the Denavit-Hartenberg approach as a rigid body tree with corresponding joint limits (Table 1).

A drilling spindle was added to form the final kinematic model of the drilling robot. The drilling spindle was configured to increase robotic drilling performance [13]. Tool center point (T^1) of the drilling robot was translated by translation vector $trans_t^1 = (0.10850, 0, 0.0598)$ and rotated by a homogenous transformation matrix defined by Euler angles with rotation sequence "$\gamma\beta\alpha$". The Euler angles for the homogeneous transformation matrix were $rot_t^1 = (0, pi/2, 0)$. The virtual model of the drilling spindle with T^1 frame is presented in Fig. 2.

Table 1. DH parameters of ABB IRB 1200 5/0.9 robot.

Joint	α_i [rad]	a_i [m]	θ_i [rad]	d_i [m]
1	$-\pi/2$	0	0	0.3991
2	0	0.448	$-\pi/2$	0
3	$-\pi/2$	0.042	0	0
4	$\pi/2$	0	0	0.451
5	$-\pi/2$	0	0	0
6	0	0	π	0.082

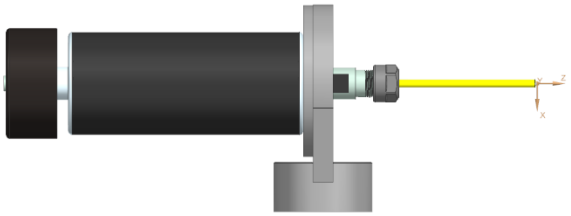


Fig. 2. Drilling spindle with $\boldsymbol{T^1}$ frame.

The work piece attached to the clamping device was added to form the final kinematic model of the service robot. Tool center point (T^2) of the service robot was translated by translation vector $trans_t^2 = (0, 0, 0.07)$, while T^2 orientation remained unchanged. The virtual assembly of the clamping device with the mounted work piece and T^2 frame is presented in Fig. 3.

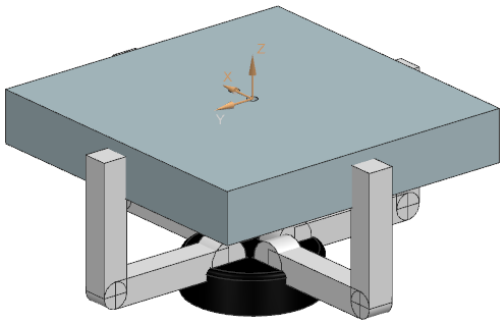


Fig. 3. Clamping device with the mounted work piece and $\boldsymbol{T^2}$ frame.

2.3. Inverse Kinematics Algorithm

The iterative Broyden-Fletcher-Goldfarb-Shanno algorithm was applied to solve inverse kinematics of the robotic drilling system. For both robots, the initial guess configuration was chosen as the robot's home position defined in Table 1. For each consecutive point along the drilling trajectory, configuration from the previous point was used as the initial guess

configuration. This ensured a smooth path without major configuration changes during task execution. A smooth path is important to achieve high geometric tolerances of the drilled hole, as for robotic drilling not only linear movement from the starting point to the end point is required, but a constant tool orientation is also required.

2.4. Genetic Algorithm

The genetic algorithm searched for the optimal placement of the service robot and the drilling trajectory relative to placement of the drilling robot, so that both robots' total cumulative manipulability along their trajectory was maximized. As manipulability is independent of the robot's first axis rotation the orientation of the service robot's base frame was not an optimization variable and was equal to the orientation of the drilling robot's base frame.

Variables:

GA variables with their lower (L. b.) and upper bounds (U. b.) are summarized in Table 2.

Table 2. System variables for GA.

Var.	L. b.	U. b.
X_t^1	0	0.901
Y_t^1	0	0.901
Z_t^1	−0.342	1.3
γ_t^2	$-\pi$	π
β_t^2	$-\pi/2$	$\pi/2$
α_t^2	$-\pi$	π
X_b^2	0.901	1.802
Y_b^2	0	0.901

Drilling robot:

The drilling robot's base frame position B_{pos}^1 was fixed to the global G frame as defined by Eq. (1).

$$B_{pos}^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (1)$$

Orientation of the drilling robot's base frame B_{eul}^1 was used as the global orientation reference.

The drilling robot's tool position coordinates T_{pos}^1 were optimization variables, defined by Eq. (2).

$$T_{pos}^1 = \begin{bmatrix} X_t^1 \\ Y_t^1 \\ Z_t^1 \end{bmatrix} \quad (2)$$

The drilling robot's tool orientation coordinates T_{eul}^1 were optimization variables, defined by Eq. (3).

$$T_{eul}^1 = \begin{bmatrix} \gamma_t^1 \\ \beta_t^1 \\ \alpha_t^1 \end{bmatrix} \quad (3)$$

Service robot:

The service robot's base position coordinates B_{pos}^2 in X- and Y-direction were optimization variables, defined by Eq. (4).

$$B_{pos}^2 = \begin{bmatrix} X_b^2 \\ Y_b^2 \\ 0 \end{bmatrix} \quad (4)$$

The service robot's work piece position coordinates T_{pos}^2 were equal to position coordinates of the drilling robot T_{pos}^1.

In general, the service robot's work piece orientation coordinates T_{eul}^2 are defined by Eq. (5).

$$T_{eul}^2 = \begin{bmatrix} \gamma_t^2 \\ \beta_t^2 \\ \alpha_t^2 \end{bmatrix} \quad (5)$$

For drilling it is important that both robots are aligned and that the drilling spindle is perpendicular to the work piece. This implies that T_{eul}^2 must be a transformation of T_{eul}^1 A convenient approach is to perform a rotation of T_{eul}^1 for pi around the local Y-axis, as expressed by Eq. (6).

$$rot_{Y_\pi} = \begin{bmatrix} cos(\pi) & 0 & -sin(\pi) \\ 0 & 1 & 0 \\ sin(\pi) & 0 & cos(\pi) \end{bmatrix} \quad (6)$$

The rotation matrix rot_{Y_π} rotates points in the XZ-plane counterclockwise through an angle π with respect to the Y-axis about the origin of a three-dimensional Cartesian coordinate system. The rotation appears counterclockwise when the axis about which it occurs points toward the observer, the coordinate system is right-handed, and the rotation angle is positive.

To obtain the rotation matrix rot_2 that represents T_{eul}^2, a transformation of T_{eul}^1 to a rotation matrix rot_1 was performed, then, rot_2 was defined by Eq. (7).

$$rot_2 = rot_{Y_\pi} * rot_1 \quad (7)$$

The transformation from T_{eul}^1 to rot_1 and the transformation from rot_2 to T_{eul}^2 both followed the same sequence of Euler angle rotations.

Variable bounds:

Only one quadrant of the system's workspace was considered in the GA optimization, as recent studies on that maximum kinematic performance of serial robots may be expected in the considered region and that the symmetry of the system eliminates the risk of excluding global extremes [14, 15].

Boundaries for positioning the drilling trajectory in X- and Y-direction were selected in the interval from 0 to the maximum reach of the robot in the corresponding direction and in Z-direction from the

minimum to the maximum reach of the robot in Z-direction.

The robot's end frame T orientations were not restricted so that the minimum and the maximum accessible Euler angles were selected as boundaries for each orientation coordinate of the robot.

The lower boundary for positioning the service robot along the X-axis was selected as the maximum reach of the robot in the corresponding direction, while the upper boundary was selected as two times the maximum reach of the robot. Boundaries in Y-axis direction were selected in the interval from 0 to the maximum reach of the robot in Y-axis direction.

Constraints:

To avoid collision postures advancing through GA generations a collision detection mechanism was included in the algorithm. Collision bodies were assigned to both robots, the drilling spindle, the clamping device, and the mounted work piece. For computational efficiency simplified geometries were used and only collisions of the fourth link on both robots, the drilling spindle, the clamping device, and the mounted work piece were considered in collision detection.

Objective function:

The ability of a robot to follow a trajectory at a prescribed velocity is depending on its kinematics. Kinematic performance of a robot can be expressed by the manipulability index [5]. For redundant robots, manipulability may be expressed by Eq. (8).

$$w = \sqrt{\det(J * J^T)}, \tag{8}$$

where J is the analytical Jacobian matrix of the robot at current configuration.

For drilling, it is important that the robot follows the drilling path smoothly and maintains the required feed rates, therefore small configuration changes of the system that are evenly distributed on the contributing axes are preferred. As drilling quality is affected by both robot's configuration equally, total manipulability w_t of the system was considered. The objective of the optimization was to minimize w_t as shown in Eq. (9),

$$w_t = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{1}{w_{i,j}}, \tag{9}$$

where $w_{i,j}$ is manipulability of each robot at its current configuration, n is the number of robots included in the system and m are trajectory points.

2.5. Drilling Trajectory

To drill the part, a tool path for drilling two through holes in normal direction to the work piece was generated. As a simplification only the drilling robot followed the drilling path by a trajectory, while the service robot kept its initial configuration determined by the GA. Therefore, only one configuration of the service robot was considered in w_t, opposed to seven configurations of the drilling robot.

The drilling tool path was generated by incremental position and orientation coordinates, shown in Table 3. Work piece frame initial position $\boldsymbol{W_{pos}}$ and orientation $\boldsymbol{W_{eul}}$ coordinates were determined by the GA and were equal to $\boldsymbol{T^1_{pos}}$ and $\boldsymbol{T^1_{eul}}$. $\boldsymbol{W_{pos}}$ was also equal to $\boldsymbol{T^2_{pos}}$ (Fig. 3).

Table 3. Incremental position coordinates along the drilling path relative to the work piece frame.

ΔX^1_t [m]	ΔY^1_t [m]	ΔZ^1_t [m]
0	0	0
0	0	$0-0.01$
0	0	$0+0.035$
0	0	$0-0.035$
$0+0.01$	0	0
0	0	$0+0.035$
0	0	$0-0.035$

The drilling tool path relative to work piece frame is presented in Fig. 4.

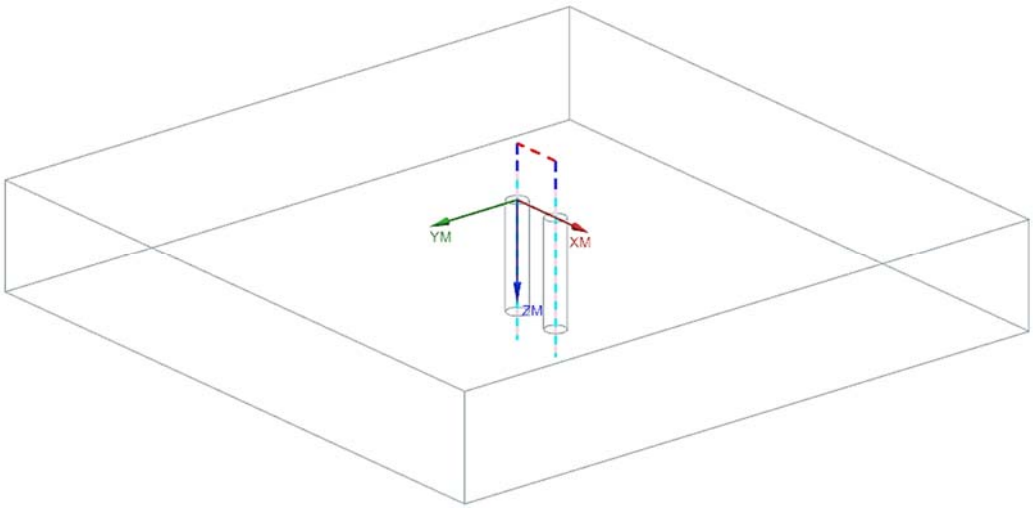


Fig. 4. Tool path for drilling two holes relative to work piece frame.

Incremental position coordinates for the service robot relative to the work piece frame $\boldsymbol{W_{pos}}$ are presented in Table 4.

Table 4. Incremental position coordinates for the service robot relative to the work piece frame.

ΔX_t^2 [m]	ΔY_t^2 [m]	ΔZ_t^2 [m]
0	0	0

The drilling robot's spindle orientation during drilling was equal to $\boldsymbol{T_{eul}^1}$ and was kept constant. Incremental orientation coordinates are presented in Table 5.

Table 5. Incremental orientation coordinates along the drilling tool path for the drilling robot.

$\Delta\gamma_t^1$ [rad]	$\Delta\beta_t^1$ [rad]	$\Delta\alpha_t^1$ [rad]
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0

Incremental orientation coordinates for the service robot T_{eul}^2 are presented in Table 6.

Table 6. Incremental orientation coordinates for the service robot.

$\Delta\gamma_t^2$ [rad]	$\Delta\beta_t^2$ [rad]	$\Delta\alpha_t^2$ [rad]
0	0	0

3. Results

The obtained optimal layout for the dual robot drilling cell found by the GA is presented in Fig. 5. The systems total manipulability w_t is 1.92e-12, whereby a value of zero would mean, that the highest possible manipulability was maintained by both robots during the complete drilling process.

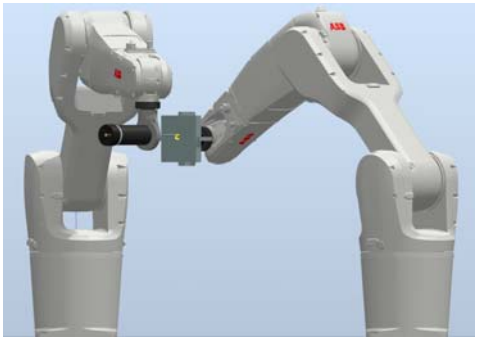


Fig. 5. Dual robot cell layout for drilling at maximum system manipulability.

The obtained optimal variable values representing the starting points and orientations of the drilling and the service robot are summarized in Table 7.

Table 7. Optimal variable values obtained by the GA.

Var	Value
X_t^1	0.58462
Y_t^1	0.26327
Z_t^1	0.50939
γ_t^2	−0.835993
β_t^2	−1.568781
α_t^2	0.103922
X_b^2	0.98253
Y_b^2	0.85038

The results are relative to the position of the drilling robot base frame (B_{pos}^1), which was fixed, according to Eq. (1). The optimal position of the service robot base frame (B_{pos}^2) is defined by Eq. (10).

$$B_{pos}^2 = \begin{bmatrix} 0.98253 \\ 0.85038 \\ 0 \end{bmatrix} \tag{10}$$

The corresponding drilling and service robot initial configurations are presented in Table 8.

Table 8. Optimal starting configuration for the drilling robot and the service robot.

Joint	Drilling robot [rad]	Service robot [rad]
1	−2.862461	−2.381730
2	−0.604760	0.754158
3	−2.704322	−0.164139
4	3.141949	−1.615860
5	1.405291	−1.635550
6	−5.272041	0.977769

The absolute trajectory position coordinates in Table 9 correspond to position coordinates of the drilling robot T^1. Position coordinates of the service robot T^2 are the first-row coordinates.

Table 9. Absolute trajectory position coordinates for the drilling (1) and the service (2) robot.

Robot	$X_t^{1,2}$ [m]	$Y_t^{1,2}$ [m]	$Z_t^{1,2}$ [m]
1 & 2	0.58462	0.26327	0.50939
1	0.59206	0.25658	0.50938
1	0.56602	0.27998	0.50943
1	0.59206	0.25658	0.50938
1	0.59208	0.25658	0.51938
1	0.56604	0.27998	0.51943
1	0.59208	0.25658	0.51938

The resulting orientation coordinates for both robots are presented in Table 10.

Table 10. Optimal tool orientation for the drilling (1) and the service (2) robot.

Robot	$\gamma_t^{1,2}$ [rad]	$\beta_t^{1,2}$ [rad]	$\alpha_t^{1,2}$ [rad]
1	−0.835993	−1.568781	0.103922
2	−2.305599	1.568781	−3.037670

4. Conclusions

This article studied the feasibility of direct robotic cooperation for robotic drilling. Optimization results show that a desired cell layout exists and that direct cooperation of robots for drilling tasks is theoretically feasible. In results, only one possible solution of a dual robot drilling cell layout is presented, but due to functional redundancy of the system, equally adapted solutions with the same minimum objective function exist. This means that additional criteria could be considered in the optimization without worsening the kinematic performance of the system. For robotic drilling, stiffness consideration could be a meaningful addition to the optimization algorithm as low stiffness is a major weakness of machining robots.

Despite its obvious advantages, direct cooperation of robots for robotic drilling was never documented before and is discussed as a novelty in this article. The most relevant advantages of direct robotic cooperation for robotic drilling include the fact that a setup with direct robotic cooperation increases the process speed, which is considered another major weakness of robotic machining also in case of shorter tool paths like drilling and deburring. Direct robotic cooperation increases system flexibility by multiplying the variety of operations that can be conducted in a relatively small space. The presented example only considered clamping and drilling, while additional operations like polishing, quality control, assembly and packing could be included to completely automatize a production process. Generally, material handling from the input to the output buffer is shortened as less actions are required to forward a work piece to its next production stage.

References

[1]. M. Hassan, D. Liu, G. Paul, collaboration of multiple autonomous industrial robots through optimal base placements, *Journal of Intelligent & Robotic Systems,* Vol. 90, Issues 1-2, 2017, pp. 113-132.

[2]. A. Verl, A. Valente, S. Melkote, C. Brecher, E. Ozturk, L. T. Tunc, Robots in machining, *CIRP Annals-Manufacturing Technology,* Vol. 68, Issue 2, 2019, pp. 799-822.

[3]. R. Koeppe, D. Engelhardt, A. Hagenauer, P. Heiligensetzer, B. Kneifel, A. Knipfer, K. Stoddard, Robot-robot and human-robot cooperation in commercial robotics applications, in *Proceedings of the Eleventh International Symposium Robotics Research*, 2005, pp. 202-216.

[4]. S. Grigorescu, S. Vatau, A. Dobra, Dedicated robot-robot cooperation, in *Proceedings of the 19th International DAAAM Symposium*, 2008, pp. 567-568.

[5]. T. Yoshikawa, Manipulability of robotic mechanisms, *The International Journal of Robotics Research,* Vol. 4, Issue 2, 1985, pp. 3-9.

[6]. G. Xiong, Y. Ding, L. M. Zhu, Stiffness-based pose optimization of an industrial robot for five-axis milling, *Robotics and Computer-Integrated Manufacturing,* Vol. 55, 2019, pp. 19-28.

[7]. J. R. P. Diaz, P. Mukherjee, A. Verl, automatic close-optimal work piece positioning for robotic manufacturing, *Procedia CIRP,* Vol. 72, 2018, pp. 277-284.

[8]. Z. Y. Lim, S. G. Ponnambalam, K. Izui, Multi-objective hybrid algorithms for layout optimization in multi-robot cellular manufacturing systems, *Knowledge-Based Systems,* Vol. 120, 2017, pp. 87-98.

[9]. G. Nicola, N. Pedrocchi, S. Mutti, P. Magnoni, M. Beschi, Optimal task positioning in multi-robot cells, using nested meta-heuristic swarm algorithms, *Procedia CIRP*, Vol. 72, 2018, pp. 386-391.

[10]. M. Farzanehkaloorazi, I. A. Bonev, L. Birglen, Simultaneous path placement and trajectory planning optimization for a redundant coordinated robotic work cell, *Mechanism and Machine Theory,* Vol. 130, 2018, pp. 346-362.

[11]. Y. Bu, W. Liao, W. Tian, L. Zhang, L. I. Dawei, Modeling and experimental investigation of Cartesian compliance characterization for drilling robot, *The International Journal of Advanced Manufacturing Technology,* Vol. 91, Issue 9, 2017, pp. 3253-3264.

[12].E. Ferreras-Higuero, E. Leal-Muñoz, J. García de Jalón, E. Chacón, A. Vizán, Robot-process precision modelling for the improvement of productivity in flexible manufacturing cells, *Robotics and Computer-Integrated Manufacturing,* Vol. 65, 2020, 101966.

[13]. Y. Lin, H. Zhao, H. Ding, Spindle configuration analysis and optimization considering the deformation in robotic machining applications, *Robotics and Computer-Integrated Manufacturing,* Vol. 54, 2018, pp. 83-95.

[14]. C. Dumas, S. Caro, M. Cherif, S. Garnier, B. Furet, Joint stiffness identification of industrial serial robots, *Robotica,* Vol. 30, 2012, pp. 649-659.

[15]. Y. Lin, H. Zhao, H. Ding, Posture optimization methodology of 6R industrial robots for machining using performance evaluation indexes, *Robotics and Computer-Integrated Manufacturing,* Vol. 48, 2017, pp. 59-72.

(004)

Recurrent Neural Network Structures for Learning Control Valve Behaviour

Camilla Sterud [1], Signe Moe [1], Mads Valentin Bram [2], Stephen Roberts [3], Jan-Peter Calliess [3]
[1] Dept. of Mathematics and Cybernetics, SINTEF Digital, Oslo, Norway
[2] Dept. of Energy Technology, Aalborg University, Esbjerg, Denmark
[3] Dept. of Engineering Science, University of Oxford, UK
E-mail: camilla.sterud@sintef.no

Summary: Control valves are ubiquitous in process control, yet they are rarely explicitly modelled. Here, we propose a theoretical valve model as a recurrent neural network (RNN) cell, allowing its parameters to be learned with gradient descent methods. Further we alter the theoretical model by incorporating a one-dimensional neural network. The models are capable of predicting valve opening from its reference value and can be easily combined with other neural network layers. We compare their performance to two long short-term memory networks (LSTMs) and showcase the performance improvements of our suggested physics-based models. In particular, we present how a gradient descent based learning algorithm finds parameters that lead to improved performance by the original theoretical valve model. The learning experiments are carried out on datasets from two different modes of operation, and we explore whether parameters that are suitable both modes can be found. The results show the benefit of using a physically inspired model for learning, like interpretable parameter values.

Keywords: Control valve, Soft sensor, Recurrent neural networks.

1. Introduction

Control valves are central and critical in the process industries as they are applied to a variety of tasks such as chemical dosing and tank level control. Predicting a valve's response to a control input can be challenging, as it is affected by nonlinear phenomena such as stiction, dead-band, backlash and hysteresis [1]. However, precise predictions are valuable, as they can be used to improve outer control loops and other estimators, leading to direct economic benefits as more efficient control systems lead to less waste and energy consumption.

In this paper, we consider prediction of the control valve opening position from its reference value, which is typically generated by an outer control loop. In particular, we suggest implementation of a newly suggested valve model, introduced in [2], as a recurrent neural network (RNN) cell, such that its parameters can be learned from valve data with gradient descent methods and be combined with other neural network (NN) layers. This alleviates the control engineer from choosing model parameters, while maintaining the interpretability of the original valve model. We also suggest a NN inspired extension of the original valve model and compare the two models' performances to two long short-term memory networks (LSTMs).

A few publications have considered modelling valve opening position, in valves experiencing stiction, using NNs [3, 4]. In [3] the authors only consider data from a simulated Choudhury model, while [4] also consider laboratory data. In contrast however, we examine both laboratory data and data from normal operation at a scaled offshore pilot plant at Aalborg University in Esbjerg [5].

2. Method

The work of [2] suggests a model for estimating the valve opening position from the opening reference signal. For illustrative purposes, this model can be visualised as an open-top cart of width $2B$, with a vertical pin sticking down into the cart – see Fig. 1. The cart will not start moving until the pin engages the cart edge and starts pulling in one direction. Thus, the pin-cart model emulates a delayed response in the cart position. The pin position is updated by a P-controller with controller gain K_p, where the control error is the difference between the cart position and a reference signal. Let x_t, m_t and r_t denote the cart position, pin position and reference position at time t, respectively. To simplify notation, we let $e_t = r_t - x_{t-1}$. Then, the pin-cart model is given by:

$$m_t = m_{t-1} + K_p e_t, \tag{1}$$

$$x_t = P_1 x_{t-1} + (1 - P_1)(m_t - P_2 B), \tag{2}$$

$$P_1 = \{1 \text{ if } |m_t - x_t| < B, 0 \text{ otherwise}\}, \tag{3}$$

$$P_2 = \text{sign}(m_t - x_{t-1}) \tag{4}$$

The pin-cart model aims to capture the behaviour of pneumatic control valves experiencing state-dependent delay effects, such as hysteresis and stiction. The reference signal r_t corresponds to the reference valve opening and the cart position represents the valve opening. The cart width B and controller gain K_p describe how responsive the valve is to changes in the reference signal.

In [2], the authors observe that the delayed step response of the valve opening depends on the initial opening position of the valve. Therefore, it is suggested to update the pin position as follows when the pin is not engaged with the cart edge:

$$m_t = m_{t-1} + \frac{\text{sign}(r_t - x_{t-1})}{a_p x_{t-1}^2 + b_p x_{t-1} + c_p}, \quad (5)$$

where a_p, b_p and c_p are free parameters. We will refer to this modified model as Pin-cart II, and the original one from Equations (1)-(4) as Pin-cart I.

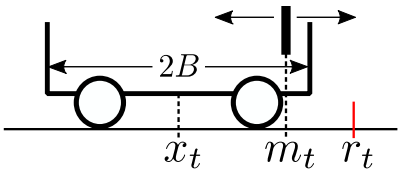


Fig. 1. Illustration of the pin-cart model.

Neural networks are popular function approximators whose representational power and generalising capabilities have been demonstrated across a wide range of applications. When trying to model dependencies in time, LSTMs are often seen as a go-to solution, at least for time-dependencies in sequences which are up to hundreds of time-steps long. LSTMs are a class of RNNs that utilise gating mechanisms to achieve both short- and long-term memory, whilst avoiding the vanishing and exploding gradient problems of vanilla RNNs. However, LSTMs are both structurally complicated, relatively hard to train and challenging to interpret, which is why there is an argument for replacing them with causal temporal convolutional networks, attention networks, transformers and residual networks [6].

In the experiments we will use two LSTMs; one with a single LSTM unit, called LSTM I, and one two-layer LSTM with four units in the first layer, and one unit in the second layer, named LSTM II.

The Pin-cart I model in Equations (1)-(4) can be implemented as a RNN cell with only two parameters, K_p and B and two hard, binary gates, P_1 and $(1 - P_1)$. We assume that measurements of the cart position are not available, so that $x_{t-1}$is the previous estimate of the cart position. In the RNN setting, x is analogous to the hidden state, and the pin position m to the cell state.

The Pin-cart II model introduces some challenges with regards to learning the parameters a_p, b_p and c_p, as (5) has one or two singularities for all sets of parameters such that $\left(a_p x_{t-1}^2 + b_p x_{t-1} + c_p\right)$ has at least one real root. One possibility is to constrain the parameter space such that the polynomial has no real roots in the interval $x_{t-1} \in [0, 1]$. However, we instead note that we might exchange (5) for a function that is more easily integrated into the NN paradigm. We suggest modelling the pin movement with the following equation:

$$m_t = m_{t-1} + \\ + \text{sign}(e_t)(a_n \tanh(b_n x_{t-1} + \text{c}_n) + d_n) \quad (6)$$

The hyperbolic tangent function (tanh) is a much-used activation function in deep learning and has a smooth, continuous derivative everywhere. We can regard this new term as a single-input single-output two-layer feed forward network with tanh activation in the first layer, and linear activation in the second layer. This version of the model will be referred to as Pin-cart III. For this model, B becomes redundant as a free parameter, and we therefore fix it to 0.1 for all models.

3. Data

Pin-cart I in Equations (1)-(4) was used to generate 20 synthetic datasets with random uniformly sampled pairs of parameters $B \in [0.01, 0.25]$ and $K_p \in [0.01, 4]$. Each simulation was run for 10 000 steps. In earlier work, two experiments were run on a Bürkert pneumatic control valve system 8802 [2]. In the first experiment the reference opening position is a series of random steps, while the second experiment shows the valve in normal, continuous operation. These three datasets will be referred to as synthetic, step and continuous, respectively.

The continuous dataset originally contains 300 000 samples. However, a greater part of the continuous dataset documents periods where the valve is in steady state at an unchanging set point. Therefore, a subset of 56 300 data points where activity is high is selected. From this subset, 29 000 are used for training, 12 500 for validation and 14 800 for testing. The step dataset consists of 120 000 measured values, where 76 800, 19 200 and 24 000 are set aside for training, validation and testing, respectively.

The data must be reshaped to be compatible with the RNNs, as these expect time sequences as input. To perform back propagation through time during training, the input length must be restricted, and becomes a hyper parameter. Denoting this input length by N_s, one input sequence starting at time $i + 1$ takes the following form: $\boldsymbol{r}_{i+1} = [r_{i+1}\, r_{i+2}\, \ldots\, r_{i+N_s}]$. Correspondingly, the prediction made based on this input is the estimated valve opening at time $i + N_s$, denoted $\hat{x}_{i+N_s}$. During training, the measured initial opening of the valve for every training example, denoted x_i, is used for intialising the cart and pin position. So while making prediction $\hat{x}_{i+N_s}$, x_i is known. During testing, only the very first measured valve opening x_0 is provided, while the rest of the predictions are only dependent on the previous predictions.

4. Neural Network Training

LSTM I, LSTM II, Pin-cart I and Pin-cart II are all trained 10 times on the training part of both the step and continuous dataset, leaving us with 80 trained

models. All models were implemented as RNNs in Python using Tensorflow.

The models were trained with the mean absolute error (MAE) loss, using the Adam optimiser and a batch size of 1024. The input sequence length was set to 500. Each model was trained until the validation loss did not improve for 50 epochs.

5. Results

Pin-cart I was trained on the synthetic dataset. demonstrates how the error in the parameters converges to zero as the number of training epochs increases.

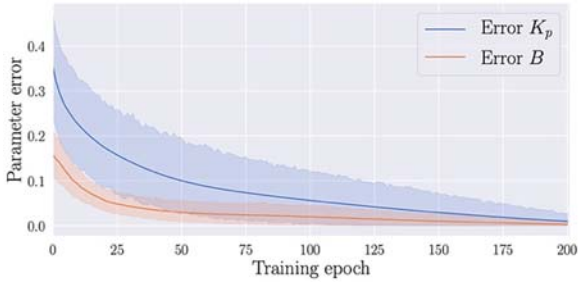


Fig. 2. The errors between the learned and true values of $\boldsymbol{B}$ and $\boldsymbol{K_p}$ when the pin-cart model is trained on the 20 synthetic datasets. The lines are the means, and the error bands represent the 95 % confidence interval.

Further, three test cases are considered:

1. The models trained on step data evaluated on the step test set;
2. The models trained on step data evaluated on the continuous test set;
3. The models trained on the unfiltered continuous data evaluated on the continuous test set.

For comparison, we include the reference signal as a baseline estimator in all tests.

5.1. Test case 1

Table 1 shows the results for Test case 1 by MAE, R^2 value and the sum of squared errors on the step test set. For this test case, LSTM II is superior to the other models by all metrics. Fig. 4 shows the performance of the models on a part of the step test set, where the blue line is the mean prediction made by the 10 models of each type, and the error band represents the standard deviation of the prediction.

5.2. Test case 2

Fig. 3 shows the expected mean absolute error (MAE) for test cases 1 and 2. In terms of MAE, LSTM II has the best performance on the step data set, followed by Pin-cart III. However, the pin-cart models experience a large reduction in MAE on the continuous datasets, in contrast to the LSTMs. Further, the pin-cart models exhibit more stable performance than the LSTMs, by having lower standard deviations.

In Table 2 we see the results from Test case 2. These numbers reveal that Pin-cart III has the best performance among the learned models on the continuous data for all metrics. However, Pin-cart III does not beat the reference baseline estimator.

Fig. 5 shows predictions by the best model of each type on parts of the continuous test set. Here we see how the pin-cart models only achieve a relatively low MAE by lagging the least behind the reference.

Table 1. Test case 1. The first number in each cell shows the mean of the 10 models of each type and the number in parenthesis shows the same metric for the best model of each type in terms of MAE.

Model	MAE	R^2	SSE
Reference	7.73 %	52.36 %	619.39
LSTM I	3.03 % (2.45 %)	93.33 % (96.61 %)	72.36 (36.37)
LSTM II	1.77 % (1.15 %)	98.12 % (99.20 %)	20.72 (9.08)
Pin-cart I	2.78 % (2.76 %)	93.26 % (93.54)	69.19 (65.97)
Pin-cart III	2.12 % (1.93 %)	96.24 % (96.70)	39.21 (35.14)

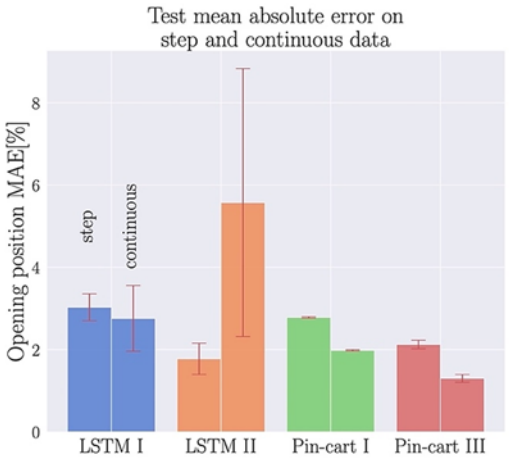


Fig. 3. Test case 2. The test MAE of each model on both the step and continuous data sets. The error bands represent the standard deviation of the MAE of the 10 models trained for each model type.

5.3. Test case 3

When trained on the continuous data, both pin-cart models perform better than the reference baseline on average. Pin-cart I produces the best performing model over all, but Pin-cart III performs slightly better on average. Neither LSTM model performs better than the baseline on average, but the best models have a performance that is on par with the average performance of the pin-cart approaches.

Fig. 6 shows predictions made by the models on part of the continuous test set. Here, the solid blue line represents the mean of the predictions made by the 10 models of each type. We see that the reference

signal is almost entirely within the error band of the predictions made by the LSTMs, while the predictions from the pin-cart models, mainly sit between the reference and the target measured valve opening.

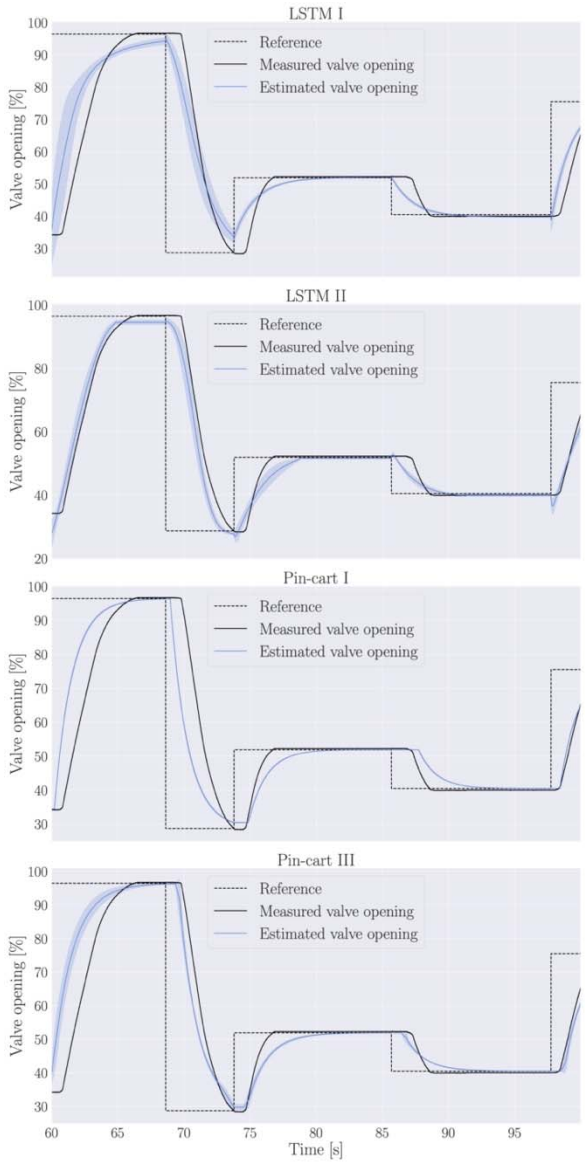


Fig. 4. Test case 1. Mean prediction on part of the step dataset by all models. The error bands indicate the one standard deviation interval of the predictions made by the 10 models of each type.

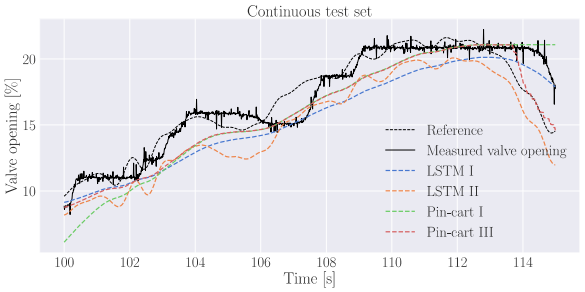


Fig. 5. Test case 2. Predictions on part of the continuous dataset by the model with the lowest MAE of each type.

Table 2. Test case 2. The first number in each cell shows the mean of the 10 models of each type and the number in parenthesis shows the same metric for the best model of each type in terms of MAE.

Model	MAE	R^2	SSE
Reference	0.95 %	99.05 %	3.97
LSTM I	4.60 % (3.41 %)	71.08 % (98.39 %)	44.45 (24.43)
LSTM II	5.64 % (2.19 %)	56.85 % (94.85 %)	89.82 (9.59)
Pin-cart I	2.23 % (2.19 %)	95.21 % (95.37 %)	13.55 (13.04)
Pin-cart III	1.39 % (1.27 %)	98.29 % (98.60 %)	4.72 (3.97)

Fig. 7 shows the mean and standard deviation of the parameters learned by the Pin-cart I models on both the step and continuous datasets. The parameters learned on the step dataset have low standard deviation, indicating that all 10 models learn similar parameters.

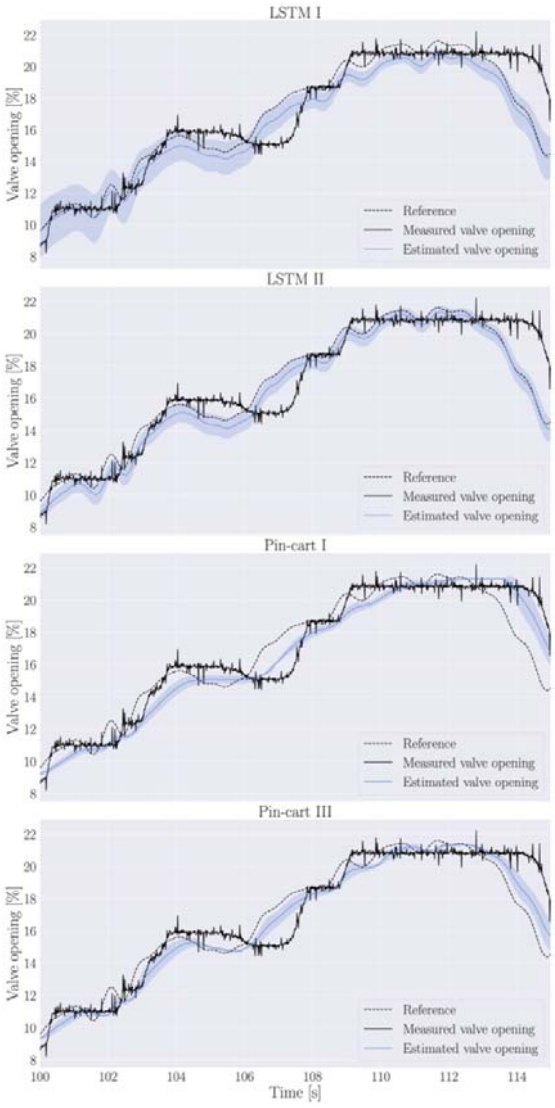


Fig. 6. Test case 3. Mean prediction on part of the continuous dataset by all models. The error bands indicate the one standard deviation interval of the predictions made by the 10 models of each type.

Table 3. Test case 3. Results on the continuous test set by the models trained on continuous data. The first number in each cell shows the mean of the 10 models of each type and the number in parenthesis presents the same metric for the best model of each type in terms of MAE.

Model	MAE	R^2	SSE
Reference	0.95 %	99.05 %	3.97
LSTM I	1.52 % (0.78 %)	95.80 % (99.40 %)	7.75 (1.43)
LSTM II	1.13 % (0.82 %)	98.62 % (99.04 %)	3.34 (2.41)
Pin-cart I	0.81 % (0.68 %)	99.23 % (99.39 %)	1.49 (1.22)
Pin-cart III	0.79 % (0.74 %)	99.18 % (99.35)	1.65 (1.30)

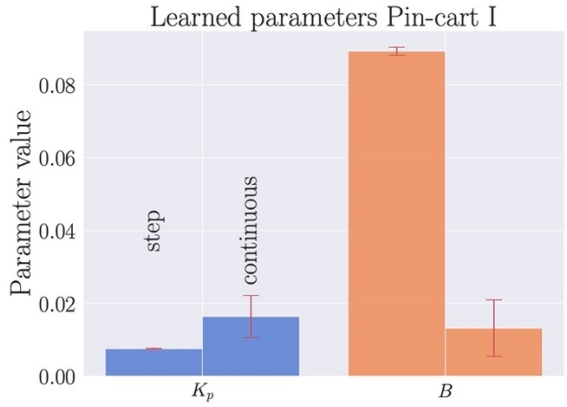


Fig. 7. Values of the learned parameters in the 20 Pin-cart I models after training. The left bar represents the parameters learned on the step data, and the right bar those learned on continuous data. The bar height represents the mean of the learned parameters of the 10 models trained on each dataset and the error bars represent one standard deviation intervals.

6. Discussion

Looking at Fig. 4, we see that the standard deviation of the predictions is virtually zero for Pin-cart I in Test case 1. By inspecting Fig. 7 we discover that K_p and B are very similar, about 0.0075 and 0.09, respectively, for all 10 models trained on the step dataset. This indicates that there exists a local minimum of the loss function with respect to B and K_p close to these values. When inspecting the parameter values of Pin-cart III, however, we do not observe such a pattern. This is reflected by the slightly larger standard deviation for the Pin-cart III predictions in Fig. 4.

In Test case 2 we observe that the pin-cart models experience a large reduction in MAE when transitioning from step to continuous data. However, as the reference baseline outperforms the learned models on all metrics, it is not expedient to use the learned models as estimators.

The fact that the pin-cart models in Test case 2 fare better than the LSTMs on the continuous datasets can be attributed to their theoretical foundation; even if their parameters are not well tuned, a minimal coherence is guaranteed by the model structure. The strong reduction in MAE can be explained by that the continuous dataset showcases smaller changes in the reference position than the step dataset.

Test case 3 demonstrates that the pin-cart models with learned parameters are well capable of describing the valve in continuous operation. The LSTMs on the other hand apparently need some luck to succeed in this, as neither outperforms the baseline on average. The learned pin-cart models perform slightly better than their non-learning counterparts with handpicked parameters when comparing to the results from [2].

One of the possible reasons it is not possible to discover parameters that are suitable for both modes of operation is that the datasets were not gathered at the same time, but about one year apart. During this time, the physical condition of the valve might have changed slightly through wear and tear.

Looking at Fig. 7 we note that, on average, K_p is twice the size and B one order or magnitude smaller for the Pin-cart I models trained on continuous data compared to the ones compared on the step data. This indicates that during continuous operation the valve is more responsive to changes in the reference than what is observed in the step dataset. Thus, even though we were not able to discover a unifying set of parameters, we are able to draw conclusions about the behaviour of the physical system by inspecting the difference in learned parameters.

Another reason for the differing parameters could also be due to the different manners in which the valve is excited by the reference signal in the two modes of operation. In the step dataset the changes are large and abrupt, whilst in the continuous dataset the reference is gradually changing. Thus, the two modes themselves are likely too different to describe with one model, as would be expected.

A last benefit of the p in-cart model as a RNN cell is the possibility to incorporate it into a larger NN structure. For instance, one might wish to use an estimate of the valve opening to estimate mass flow through the valve. In such a case, the pin-cart model can be easily integrated and possibly updated in an end-to-end fashion.

7. Conclusion

In this paper, we showed that the parameters of a previously proposed valve model can be learned using a gradient descent method by implementing the model as a RNN cell in Python using Tensorflow. It was suggested to alter the model by integrating a one-dimensional NN, which improves the performance in two out of three test cases.

Three different datasets, two of which were gathered during two different modes of operation by the same valve, were considered. A single set of parameters that could satisfactorily explain the two modes could not be found. However, two distinct sets yielding good performance were discovered. As the

learning models are based on a physical model, we were able to interpret the difference in valve behaviour based on the learned parameters.

The suggested models were compared to two LSTMs, and while one of the LSTMs outcompete the suggested models in one test case, it does not have the benefit of having physically interpretable parameters.

In addition, as the suggested models are implemented as RNN cells, they can be incorporated into larger NNs to form more complex estimators, for instance for predicting mass flow.

Acknowledgements

This work was supported by the industry partners Borregaard, Elkem, Hydro, Yara and the Research Council of Norway through the project TAPI: Towards Autonomy in Process Industries, project number 294544. S. Roberts and J. Calliess are grateful to the Oxford-Man Institute and S. Roberts to the UK Royal Academy of Engineering.

References

[1]. Fisher Controls International LLC Emerson Automation Solutions, https://www.emerson.com/documents/automation/control-valve-handbook-en-3661206.pdf

[2]. M. Bram, J. Calliess, S. Roberts, D. S. Hansen, Z. Yang, Analysis and modeling of state-dependent delay in control valves, in *Proceedings of the International Federation of Automatic Control World Congress*, Berlin, 2020.

[3]. H. Zabiri, N. Mazuki, A black-box approach in modeling valve stiction, *International Journal of Mechanical and Mechatronics Engineering*, Vol. 4, Issue 8, 2010, pp. 605-612.

[4]. S. Sharma, V. Kumar, K. Rana, Pneumatic control valve stiction modeling using artificial neural network, in *Proceedings of the International Conference on Emerging Trends in Computing and Communication Technologies*, Dehradun, India, 2017.

[5]. M. Bram, S. Jespersen, D. S. Hansen, Z. Yang, Control-oriented modeling and experimental validation of a deoiling hydrocyclone system, *Processes*, Vol. 8, 2020, 1010.

[6]. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, in *Proceedings of the 31st International Conference on Neural Information Processing (NIPS'17)*, Los Angeles, 2017, pp. 6000-6010.

(005)

Velocity Planning of a Robotic Task Enhanced by Fuzzy Logic and Dynamic Movement Primitives

B. Maggioni [1], E. Marescotti [1], A. M. Zanchettin [1], D. Piga [2], L. Roveda [2]
[1] Politecnico di Milano, Milano, Italy
[2] Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Scuola Universitaria Professionale della Svizzera Italiana (SUPSI), Università della Svizzera Italiana (USI) IDSIA-SUPSI, Manno, Switzerland
E-mails: beatrice.maggioni@mail.polimi.it, elia.marescotti@mail.polimi.it, loris.roveda@idsia.ch

Summary: Many industrial tasks, such as welding and sealing, require not only a precise path reference, but also an advanced velocity planning in order to achieve the target quality for the final products. In this paper, a novel approach is proposed to perform robotic trajectory planning. The developed algorithm exploits Fuzzy Logic (FL) to relate the path features (such as curves or sharp edges) to the proper execution velocity. Such a computed velocity reference is then used as an input for Dynamical Movement Primitives (DMP), providing the reference signals to the robot controller. The main improved methodology features are: path-based velocity planning, extension of DMP to variable velocity reference and smoothing of the velocity reference including robot velocity/acceleration limits. The algorithm can be implemented in a collaborative framework, defining a compliant controller embedded into the DMP for online trajectory planning.

Keywords: Autonomous robotics, Collaborative robotics, DMP, Fuzzy Logic, Trajectory planning.

1. Introduction

Within Industry 4.0 paradigm, industrial tasks are re-designed enhancing the automatization of the production lines. In such a context, robotics plays a fundamental role, in particular considering the human-centered solutions that can be implemented [1].

To relieve the operators from tedious and hard coding of each specific application, robots must be able to learn and perform a reference task, exploiting autonomous planners for motion generation. Such topic is critical in many applications, like sealing and welding [2, 3], where trajectory planning and optimization is the main objective [4, 5]. The aim is, therefore, to automatically assess high-accuracy performance in trajectory tracking to achieve the target task quality.

1.1. Related Works

Trajectory planning is a hot-research topic. In [6], a widely used algorithm for welding applications is described. The planner finds the optimized motion for both the robot end-effector and joints of a welding robots, but it doesn't set the velocity along the path. In [7], a sealing task is performed using global planning interpolation and trapezoidal speed profile, but without considering any coupling between the involved degrees of freedom and without a variable velocity. In [8], Dynamical Movement Primitives (DMP) are assessed for movement sequencing trajectory planning employing velocity continuity between blocks, but there is not a punctual characterization of the velocity in the single block.

Commonly, the traditional approaches for motion planning [9] do not address the problem of the punctual characterization along the path's natural coordinate. Indeed, such approaches optimize the motion reducing the execution time, but these procedures do not take into account the execution path geometry. Those algorithms work really well in open space movements, while failing in material deposition tasks in which it is fundamental to precisely define the optimal time with a direct correlation to the optimal quality of the final result [10]. The aim of the proposed work is to reduce the total task time by automatically imposing a proper execution velocity along the path natural coordinate (*i.e.*, considering the path geometry).

1.2. Paper Contribution

Taking as a reference an automatic sealing task (within H2020 CS2 ASSASSINN project), the here presented contribution aims to design a trajectory planner able to generate the robot's reference motion to properly manage the sealant deposition.

The task execution velocity, which strongly affects the material deposition, is the main design and control parameter. The velocity reference has to be managed considering the deposition path, taking into account its geometrical features (such as sharp edges, curves, etc.) to avoid a surplus/shortage of sealing material during the deposition, while smoothing out vibrations [7]. To do that, the trajectory planning problem must consider both geometrical path features and hardware limitations (robot velocity/acceleration limits).

The here presented paper proposes an adaptive path-based task execution velocity, with a combination of Fuzzy Logic (FL) and DMP methods for the path velocity definition and for the generation of the approximating smoothed trajectory.

The FL relates the path features to the proper execution velocity, ensuring a proper sealant extrusion

quality. The computed velocity is then used as input for the DMP, so that it is possible to generate the reference velocity for the robot controller [10].

While the FL methodology has been selected due to its capabilities in experimental I/O mapping [11], DMP were selected for their capabilities in trajectories representation and time/space scaling [12].

Simulations and experiments have been performed by means of a custom setup, depicted in Fig. 1. Achieved results highlight the trajectory planning capabilities of the proposed framework, considering a complex reference path. A comparison with standard DMP (*i.e.*, defining an almost constant velocity reference along the path) has been performed.

Fig. 1. Experimental setup: Franka Emika Panda robot with a Makita caulking gun connected through a custom flange.

2. Methodology

The algorithm structure of the proposed approach, depicted in Fig. 2, is composed by trajectory pre-processing steps, where FL is exploited, and by the trajectory execution step in which the DMP are carried out using a novel modified input approach.

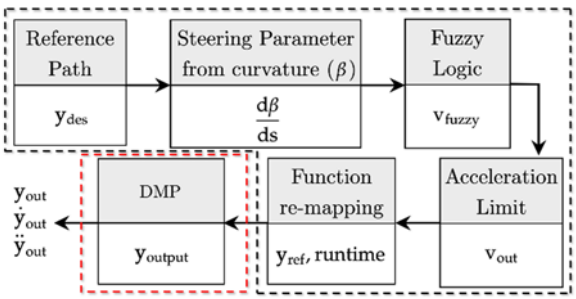


Fig. 2. Trajectory planning framework. Pre-processing block is highlighted in **B**. Execution block is highlighted in **R**.

2.1. Pre-processing Analysis

The algorithm firstly re-samples the path, $\mathbf{y}_{des}$, in order to have all the teaching points equally spaced along the path natural coordinate.

Then, the automatic path recognition is performed, through the generation and analysis of a parameter called "*steer*", which defines the local path curvature. It is defined as follows (Fig. 3):

- Vector $\mathbf{v}_{i,i+n}$ is defined, connecting the considered path point with another one *n* steps forward;
- Vector $\mathbf{v}_{i+1,i+1+n}$ is defined, connecting the *next* point with the *n+1* steps forward point;
- Angles $\beta(s_i)$ and $\beta(s_{i+1})$ are computed (between the horizontal axis and $\mathbf{v}_{i,i+n}$ / $\mathbf{v}_{i+1,i+1+n}$, respectively);
- *Steer* is defined as: $steer_i = \beta(s_{i+1}) - \beta(s_i)$.

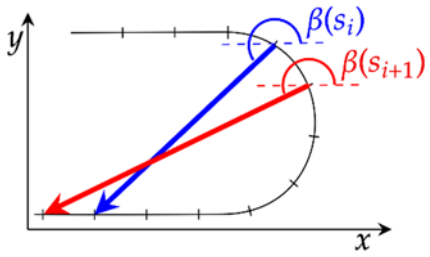


Fig. 3. Example of vector definition during a path curve. In the example: n = 7.

The absolute value of the *steer* parameter is then used as an input to the FL: each value locally describes a certain path feature (*i.e.* straight lines have *steer = 0*). The FL relates, therefore, the path geometry to the reference velocity through the generation of an experimental I/O non-linear law (Fig. 4). The velocity imposed by the FL is the maximum one at which a critical path's feature can be executed with proper sealant extrusion quality.

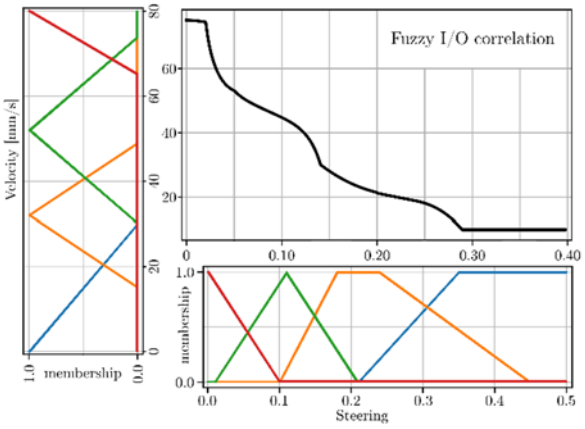


Fig. 4. FL input (*steer*) and output (*velocity*) shape functions. The Fuzzy rules are represented by the colors (*i.e.*, low *steer*, **R**, corresponds to high *velocity* and vice versa). In **B** the I/O non-linear correlation is reported.

Entering the acceleration limit block (Fig. 2), the planned trajectory in output from the FL is corrected accordingly to the end-effector linear acceleration/deceleration limits, by relaxing the time intervals at which each spatial point is reached. The

two cases, acceleration and deceleration overshoots, are depicted in Fig. 5:

- *if* $a(s_{i+1}) > a_{max}$ *then* set $a(s_{i+1}) = a_{max}$, such that $v_{out}(s_{i+1}) < v_{fuzzy}(s_{i+1})$ and the new time instant is longer;
- *if* $a(s_{i+1}) < -a_{max}$ *then* fix $v_{out,back}(s_{i+1}) = v_{fuzzy}(s_{i+1})$, and lower the previous computed velocity values, $v_{out,old}(s_i, s_{i-1}, \ldots)$, up to the convergence of related deceleration values (backward).

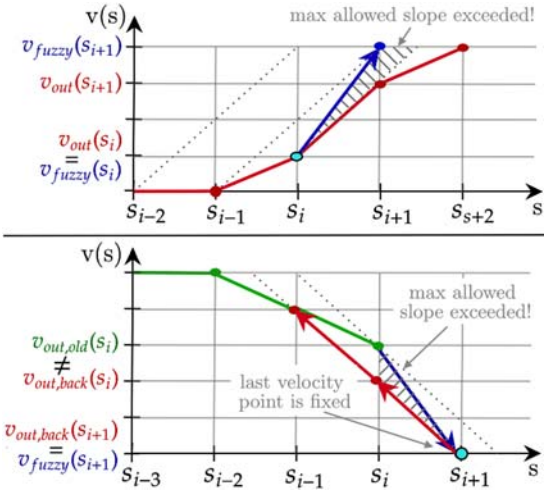


Fig. 5. Time-shift for the acceleration limit: in (**B**&**G**) the computed *v(s)* that exceeds the limit, while in (**R**) the new (feasible) velocity reference.

It must be noticed that in the deceleration case, $v_{out,back}(s_{i+1})$ cannot be increased to reduce the deceleration (as done for the acceleration), so that the maximum punctual velocity does not exceed the one computed by the FL controller.

2.2. Path Execution

The computed trajectory $\mathbf{y}_{ref}$ is finally converted to be a time function, rather than a natural coordinate used in the pre-processing stages of the algorithm. Then, it is used as an input to the DMP framework, which approximates the path with smoothing features and providing continuity to the velocity.

The novel contribution proposed by this paper is related to the capabilities of the modified DMP to make use of the input path (which has to be reproduced) to achieve a final task execution with a punctual velocity characterization along the path natural coordinate (*i.e.*, adapting the execution velocity).

The provided algorithm can be executed offline, just by sending the reference positions to the robot controller, but it also allows to define an on-line human-robot collaborative framework by embedding a compliant controller to manage external interactions sending real-time signals to the robot and to perform trajectory error recovery [13]. The trajectory $\mathbf{y}_{out}$ is finally fed into the robot position controller commanding the Cartesian end-effector signals using the built-in C++ library, *lib_franka*.

3. Results

The proposed framework has been successfully tested both in numerical simulations with a Python code and experimentally on a Franka Emika Panda robot (Fig. 1).

The numerical analysis focuses on the approximation of a taught path, which has been executed with both standard and modified DMP approaches (Fig. 6). Considering a proper tuning of the DMP, it is possible to achieve in both cases a proper trajectory reproduction. However, considering the standard DMP, the acceleration has unwanted sharp peaks which would cause vibrations in the real experiments. Moreover, in this case the total task time must be fixed a priori, without an optimization with respect to the path length and geometry. Conversely, the novel modified DMP approach permits to control the velocity along all the path, and consequently to assign an optimized total runtime as function of the path complexity. To compare the results, the standard runtime has been set to be equal to the optimized time of the modified input approach (resulting in a different velocity profile).

In Fig. 6 it is possible to see that the testing path (Fig. 6a) is performed at constant velocity if the standard DMP approach is considered (Fig. 6b), and with modulated velocity using the modified DMP approach (Fig. 6c): the straight lines (orange) are executed at a higher velocity with respect to the small curves (light blue).

Several testing paths have been studied in order to check the consistency of the algorithm. In Fig. 7 and in Fig. 8 a complete generic path which has different complex geometries (small radius curves, sharp edges and a saw tooth profile) is reported. The experimental executions show the sealant extrusion comparison between standard DMP approach (Fig. 7) and the modified DMP approach (Fig. 8). The standard execution shows large vibrations in correspondence of the sharp directional changes (*i.e.*, in the saw tooth) and the overall extrusion is not uniform, reducing the final quality. Instead, the result of the modified DMP approach smooths out most of the vibrations, achieving a higher quality deposition along the path.

4. Conclusions

The presented paper proposes a framework for autonomous trajectory planning, being able to take into account geometrical path features for velocity definition, while also considering velocity/ acceleration limits. The planned trajectory shows an execution velocity that is dependent on the geometrical path characteristics, making possible to reduce the total execution time, while obtaining the target quality for the specific deposition task.

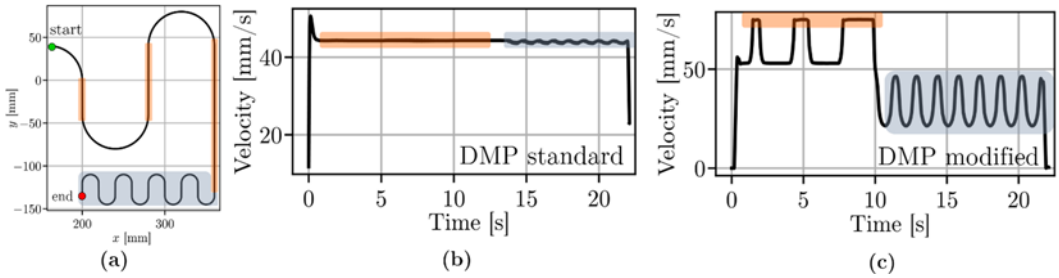


Fig. 6. Two different executions of the path shown in (a): in (b) the output velocity from a classical DMP formulation is presented, while in (c) the newly proposed method permits a continuous modulation of the velocity reference.

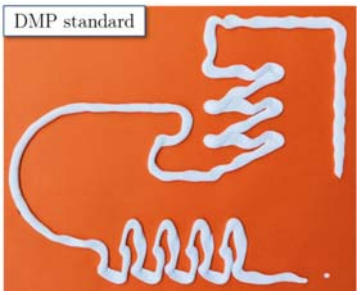


Fig. 7. Sealant extrusion with standard DMP approach.

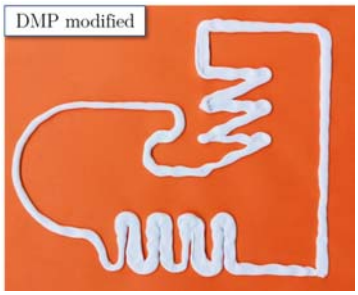


Fig. 8. Sealant extrusion with modified DMP approach.

The presented algorithm needs a manual tuning of the FL controller in order to characterize the velocity shape functions referred to some reference geometries. Current work is devoted to optimize such shape functions using a pairwise preference-based algorithm, which exploits operator's judgements [14].

While the here presented trajectory planning is performed offline, the proposed DMP framework is under implementation for real-time trajectory planning, embedding a compliant controller into the DMP framework to manage external interaction (*e.g.*, for human-robot collaboration purposes and disturbance recovery).

Acknowledgements

The work has been developed within the project ASSASSINN, funded from H2020 CleanSky 2 under grant agreement n. 886977.

References

[1]. R. Galin, R. Meshcheryakov, Automation and robotics in the context of Industry 4.0: The shift to collaborative robots, *IOP Conference Series: Materials Science and Engineering*, Vol. 537, 2019, 032073.

[2]. J. Ogbemhe, K. Mpofu, Towards achieving a fully intelligent robotic arc welding: A review, *Industrial Robot: An International Journal*, Vol. 42, Issue 5, 2015, pp. 475-484.

[3]. Z. Chen, et al., An optimized trajectory planning for welding robot, *Materials Science and Engineering*, Vol. 324, 2018, 012009.

[4]. L. Biagiotti, C. Melchiorri, Trajectory Planning for Automatic Machines and Robots, *Springer Science & Business Media*, 2008.

[5]. Y. Chen, L. Roveda, D. J. Braun, Efficiently computable constrained optimal feedback controllers, *IEEE Robotics and Automation Letters*, Vol. 4, Issue 1, 2018, pp. 121-128.

[6]. J. De Maeyer, et al., Cartesian path planning for Arc welding robots: Evaluation of the Descartes algorithm, in *Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'17)*, 2017, pp. 1-8.

[7]. L. Anderlucci, Smooth trajectory planning for anthropomorphic industrial robots employed in continuous processes, PhD Thesis, *Politecnico di Torino*, 2019.

[8]. T. Kulvicius, et al., Joining movement sequences: modified dynamic movement primitives for robotics applications exemplified on handwriting, *IEEE Transactions on Robotics*, Vol. 28, Issue 1, 2011, pp. 145-157.

[9]. B. Siciliano, et al., Robotics: Modelling, Planning and Control, *Springer Science & Business Media*, 2010.

[10]. P. Pastor, et al., Learning and generalization of motor skills by learning from demonstration, in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2009, pp. 763-768.

[11]. T. J. Ross, et al., Fuzzy Logic with Engineering Applications, *John Wiley & Sons Ltd*, 2004.

[12]. A. J. Ijspeert, et al., DMP: Learning attractor models for motor behaviors, Neural Computation, Vol. 25, 2013, pp. 328-373.

[13]. M. Karlsson, et al., Two DoFs control for trajectory tracking and perturbation recovery during execution of dynamical movement primitives, *IFAC-PapersOnLine*, Vol. 50, Issue 1, 2017, pp. 1923-1930.

[14]. A. Bemporad, D. Piga, Active preference learning based on radial basis functions, *arXiv Preprint*, arXiv:1909.13049, 2019.

(006)

Sufficient Conditions for the Existence of Periodic Solutions to a Modified Elman Neural Network

Z. Kovacheva [1,2] and V. Covachev [1]
[1] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria
[2] University of Mining and Geology, Sofia, Bulgaria
Tel.: + 359885018822
E-mail: zkovacheva@hotmail.com

Summary: Elman neural network is a recurrent neural network. Compared with traditional neural networks, an Elman neural network has additional inputs from the hidden layer, which form a new layer called the context layer. Elman neural network can be applied to solve prediction problems of discrete time sequences. In the present contribution, for a modified Elman neural network with a periodic input, we present sufficient conditions for the existence of a periodic output by using Mawhin's continuation theorem of the coincidence degree theory. An example is given of an Elman neural network satisfying these sufficient conditions. A periodic solution is found for a particular choice of the weights and periodic inputs.

Keywords: Elman neural network, Hidden layer, Context layer, Periodic input and output, Mawhin's continuation theorem.

1. Introduction

Artificial neural networks are computational paradigms which implement simplified models of their biological counterparts, biological neural networks. Biological neural networks are the local assemblages of neurons and their dendritic connections that form the human brain. Neural networks process information in a similar way the human brain does. The network is composed of a large number of highly interconnected processing elements (neurons) working in parallel to solve a specific problem. Neural networks learn by example.

Elman neural network [1] is a kind of recurrent neural networks. Compared with traditional neural networks, an Elman neural network has additional inputs from the hidden layer, which form a new layer called the context layer. So the standard back-propagation algorithm used in Elman neural network is called Elman back-propagation algorithm. Elman neural network can be applied to solve prediction problems of discrete time sequences [2-4].

The Elman neural network is one of the most widely used and most effective neural network models in artificial neural networks and has powerful processing ability for nonlinear decisions [5, 6]. The Elman neural network can be considered as a special kind of feedforward neural network with additional memory neurons and local feedback. Because of its better learning efficiency, approximation ability, and memory ability than other neural networks, the Elman neural network can not only be used in time series prediction, but also in system identification and prediction [7, 8, 4, 9].

In our recent paper [10], for a modified Elman neural network with a periodic input, we obtained sufficient conditions for the existence of a periodic output by using Mawhin's continuation theorem of the coincidence degree theory. In the present contribution, an example is given of an Elman neural network satisfying these sufficient conditions. A periodic solution is found for a particular choice of the weights and periodic inputs.

2. Main Results

We consider a modified Elman network with r nodes in the input layer, n nodes in the hidden and context layers, respectively, and m nodes in the output layer which adds a self-feedback factor α, $0 < \alpha < 1$, in the context nodes, based on the traditional Elman neural network [4, 11]. Its mathematical model is:

$$x_i(k) = f_i\left(\sum_{j=1}^{n} a_{ij} x_j^C(k) + \sum_{j=1}^{r} b_{ij} u_j(k-1)\right), \qquad (1)$$

$$x_i^C(k) = \alpha x_i^C(k-1) + x_i(k-1),\ i = \overline{1,n}, \qquad (2)$$

$$y_i(k) = g_i\left(\sum_{j=1}^{n} c_{ij} x_j(k)\right),\ i = \overline{1,m}, \qquad (3)$$

for $k \in \mathbb{N}$. Here u is the input, x is the output of the hidden layer, x^C is the output of the context layer, and y is the output of the output layer. The matrices with entries a_{ij}, b_{ij}, and c_{ij} are the weights respectively of the context nodes, input nodes and hidden nodes; f_i are the transfer functions of the hidden layer often taken as sigmoid functions, g_i are the transfer functions of the output layer and are often taken as linear functions.

Clearly, for a given input $u(k)$, $k \in \{0\} \cup \mathbb{N}$, and initial values $x(0), x^C(0)$ we can find the output $y(k), k \in \mathbb{N}$, from system (1)-(3).

Now suppose that the input $u(k)$ is N-periodic for some $N \in \mathbb{N}$, that is, $u(k+N) = u(k)$, $k \in \{0\} \cup \mathbb{N}$. We are looking for an N-periodic output $y(k)$, $k \in \mathbb{N}$. This means that for a suitable choice of the initial values $x(0), x^C(0)$ the output $y(k)$ is N-periodic. For this purpose, it suffices that the output $x(k)$ of the hidden layer is N-periodic.

We make the following assumptions:

A1. There exist positive constants L_i, $i = \overline{1,n}$, such that

$$|f_i(x_i) - f_i(\tilde{x}_i)| \le L_i|x_i - \tilde{x}_i|,$$

for all $x_i, \tilde{x}_i \in \mathbb{R}, i = \overline{1,n}$.

A2. The following inequality holds:

$$\min_{i=\overline{1,n}} \left(1 - \frac{1}{1-\alpha}\sum_{j=1}^{n} L_j|a_{ji}|\right) > 0$$

In order to formulate our last assumption, we introduce the $n \times n$-matrix

$$\mathcal{A} = \left(\delta_{ij} - \frac{L_i}{1-\alpha}|a_{ij}|, i,j = \overline{1,n}\right),$$

where δ_{ij} is the Kronecker delta, and assume that

A3. The matrix $\mathcal{A}$ is an *M*-matrix.

Assumption **A3** implies that the matrix $\mathcal{A}$ is nonsingular and its inverse has only nonnegative entries [12, 13].

The main result of [10] is the following theorem.

Theorem 1. *Suppose that assumptions **A1-A3** hold. Then the system of equations (1), (2) has at least one N-periodic solution $x(k)$.*

Theorem 1 is proved using Mawhin's continuation theorem of coincidence degree theory [14].

3. Example

Consider a modified Elman neural network with $r = 2, n = 3$ and $m = 4$ (see Fig. 1 at the end of the paper). Suppose that $\alpha = \frac{1}{2}$, the transfer functions $f_i(x)$, $i = \overline{1,3}$, of the hidden layer all equal the sigmoid function $f(x) = \frac{1}{1+e^{-x}}$, u_1, u_2 are arbitrary N-periodic functions for some positive integer N, say,

$$u_1(k) = k - \left[\frac{k}{N}\right]N, u_2(k) = k + 1 - \left[\frac{k}{N}\right]N,$$

$k \in \{0\} \cup \mathbb{N}$, where $[v]$ is the greatest integer in the real number v, that is,

$$u_1(0) = 0, u_1(1) = 1, \dots, u_1(N-1) = N-1,$$

$$u_1(N) = 0,\ u_1(N+1) = 1, \dots, u_1(2N-1) = N-1, u_1(2N) = 0, u_1(2N+1) = 1, \dots,$$

$$u_2(N) = 1,\ u_2(N+1) = 2, \dots, u_2(2N-1) = N, u_2(2N) = 1, u_2(2N+1) = 2, \dots$$

a) Further on, let us assume, for the sake of simplicity, that the weights a_{ij}, $i,j = \overline{1,3}$, of the context nodes all equal $\frac{1}{2}$, the transfer functions g_i, $i = \overline{1,4}$, the weights b_{ij}, $i = \overline{1,3}$, $j = \overline{1,2}$, of the input nodes and c_{ij}, $i = \overline{1,4}$, $j = \overline{1,3}$, of the hidden nodes are arbitrary.

Then, assumption **A1** is satisfied with $L_i = \frac{1}{4}$, $i = \overline{1,3}$, assumption **A2** is also satisfied since

$$1 - \frac{1}{1-\alpha}\sum_{j=1}^{3} L_j|a_{ji}| = \frac{1}{4}, i = \overline{1,3}$$

Finally, the matrix

$$\mathcal{A} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

is an *M*-matrix with inverse

$$\mathcal{A}^{-1} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Since all assumptions of Theorem 1 is satisfied, the modified Elman neural network under consideration has an N-periodic output $y(k)$, $k \in \mathbb{N}$.

Let us assume that, moreover, $b_{ij} = \frac{1}{2}$, $i = \overline{1,3}$, $j = \overline{1,2}$, and $N = 3$. Then, the system of equations (1), (2) takes the form

$$x_i(k) = f\left(\frac{1}{2}\left(\sum_{j=1}^{3} x_j^C(k) + \sum_{j=1}^{2} u_j(k-1)\right)\right), \quad (4)$$

$$x_i^C(k) = \frac{1}{2}x_i^C(k-1) + x_i(k-1), \quad k \in \mathbb{N}, \quad i = \overline{1,3}, \quad (5)$$

where

$$u_1(k) = \begin{cases} 1, & k \equiv 1 \pmod 3 \\ 2, & k \equiv 2 \pmod 3, \\ 0, & k \equiv 0 \pmod 3 \end{cases}$$

$$u_2(k) = \begin{cases} 2, & k \equiv 1 \pmod 3 \\ 3, & k \equiv 2 \pmod 3, \\ 1, & k \equiv 0 \pmod 3 \end{cases}$$

It suffices to find the initial values $x_i(0)$, $x_i^C(0)$, $i = \overline{1,3}$, so that

$$x_i^C(3) = x_i^C(0), x_i(3) = x_i(0), i = \overline{1,3} \quad (6)$$

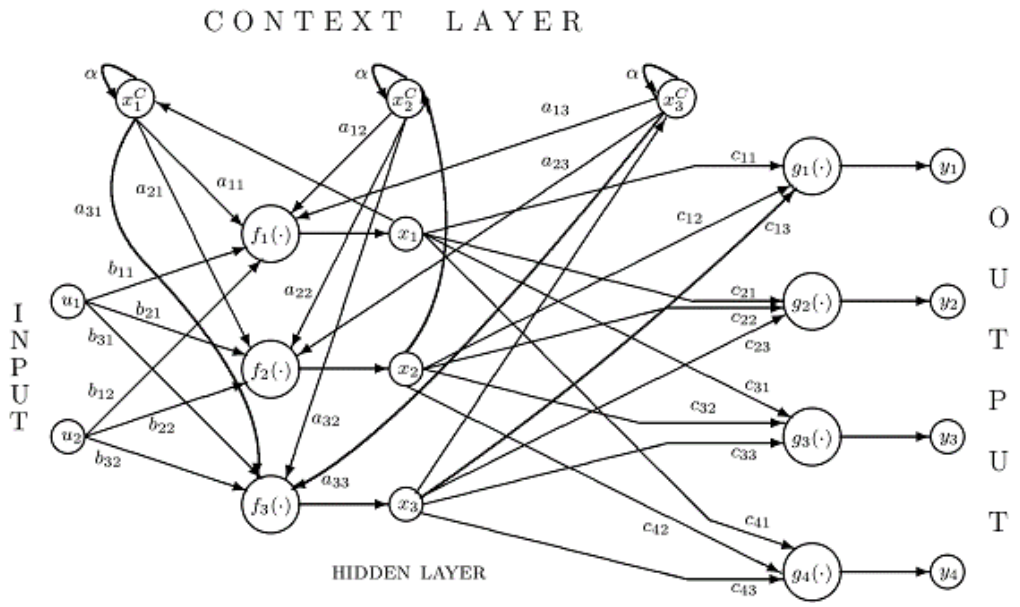


Fig. 1. An Elman neural network with $r = 2, n = 3$ and $m = 4$.

Equations (4), (5) imply that $x_1(k) = x_2(k) = x_3(k)$, $k \in \mathbb{N}$, and $x_1^C(k) = x_2^C(k) = x_3^C(k)$, $k = 2,3,...$ Thus, in order to satisfy equations (6), the initial conditions must be chosen so that $x_1(0) = x_2(0) = x_3(0)$ and $x_1^C(0) = x_2^C(0) = x_3^C(0)$. System (4), (5) reduces to

$$x(k) = f\left(\frac{1}{2}\begin{pmatrix}3x^C(k) + u_1(k-1) + \\ +u_2(k-1)\end{pmatrix}\right), \quad (7)$$

$$x^C(k) = \frac{1}{2}x^C(k-1) + x(k-1), k \in \mathbb{N} \quad (8)$$

We have found that the initial values $x^C(0), x(0)$ satisfying (6) are (approximately) $x^C(0) = 1.9634$, $x(0) = 0.9810$. The first 4 values of the 3-periodic solution of system (7), (8) are presented in Table 1.

Table 1. A 3-periodic solution of system (7), (8).

	$k=0$	$k=1$	$k=2$	$k=3$
$x^C(k)$	1.9634	1.9627	1.9504032	1.9633761
$x(k)$	0.9810	0.9690532	0.9881745	0.9810172

b) Next, let us assume that $a_{11} = \frac{1}{2}$, $a_{12} = a_{22} = \frac{1}{4}$, $a_{13} = a_{21} = a_{33} = \frac{1}{8}$, $a_{23} = a_{31} = \frac{1}{16}$, $a_{32} = \frac{1}{32}$, the transfer functions $g_i, i = \overline{1,4}$, the weights $b_{ij}, i = \overline{1,3}, j = \overline{1,2}$, of the input nodes and $c_{ij}, i = \overline{1,4}, j = \overline{1,3}$, of the hidden nodes are still arbitrary.

Then, assumption **A1** is still satisfied with $L_i = \frac{1}{4}$, $i = \overline{1,3}$, assumption **A2** is also satisfied since

$$1 - \frac{1}{1-\alpha}\sum_{j=1}^{3} L_j|a_{j1}| = \frac{5}{16},$$
$$1 - \frac{1}{1-\alpha}\sum_{j=1}^{3} L_j|a_{j2}| = \frac{47}{64},$$
$$1 - \frac{1}{1-\alpha}\sum_{j=1}^{3} L_j|a_{j3}| = \frac{27}{32}$$

Finally, the matrix

$$\mathcal{A} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{8} & -\frac{1}{16} \\ -\frac{1}{16} & \frac{7}{8} & -\frac{1}{32} \\ -\frac{1}{32} & -\frac{1}{64} & \frac{15}{16} \end{pmatrix}$$

is an *M*-matrix with inverse

$$\mathcal{A}^{-1} = \begin{pmatrix} 1.3536229 & 0.1951023 & 0.0967449 \\ 0.0983574 & 1.1577144 & 0.0451476 \\ 0.0467601 & 0.0257986 & 1.070644 \end{pmatrix}$$

Since all assumptions of Theorem 1 are satisfied, the modified Elman neural network under consideration has an N-periodic output $y(k), k \in \mathbb{N}$.

Let us assume that, moreover, $b_{11} = \frac{1}{2}$, $b_{12} = b_{21} = \frac{1}{3}$, $b_{22} = b_{31} = \frac{1}{4}$, $b_{32} = \frac{1}{8}$, and $N = 3$. Then, the system of equations (1), (2) takes the form

$$x_i^C(k) = \frac{1}{2}x_i^C(k-1) + x_i(k-1), k \in \mathbb{N}, i = \overline{1,3},$$

$$\left.\begin{aligned} x_1(k) &= f\left(\frac{1}{2}x_1^C(k) + \frac{1}{4}x_2^C(k) + \frac{1}{8}x_3^C(k) + \frac{1}{3}\right) \\ x_2(k) &= f\left(\frac{1}{8}x_1^C(k) + \frac{1}{4}x_2^C(k) + \frac{1}{16}x_3^C(k) + \frac{1}{4}\right) \\ x_3(k) &= f\left(\frac{1}{16}x_1^C(k) + \frac{1}{32}x_2^C(k) + \frac{1}{8}x_3^C(k) + \frac{1}{8}\right) \end{aligned}\right\}$$
$$k \equiv 1 \pmod 3,$$

$$\left.\begin{aligned} x_1(k) &= f\left(\frac{1}{2}x_1^C(k)+\frac{1}{4}x_2^C(k)+\frac{1}{8}x_3^C(k)+\frac{7}{6}\right)\\ x_2(k) &= f\left(\frac{1}{8}x_1^C(k)+\frac{1}{4}x_2^C(k)+\frac{1}{16}x_3^C(k)+\frac{5}{6}\right)\\ x_3(k) &= f\left(\frac{1}{16}x_1^C(k)+\frac{1}{32}x_2^C(k)+\frac{1}{8}x_3^C(k)+\frac{7}{8}\right)\end{aligned}\right\}$$
$$k \equiv 2 \pmod 3,$$

$$\left.\begin{aligned} x_1(k) &= f\left(\frac{1}{2}x_1^C(k)+\frac{1}{4}x_2^C(k)+\frac{1}{8}x_3^C(k)+2\right)\\ x_2(k) &= f\left(\frac{1}{8}x_1^C(k)+\frac{1}{4}x_2^C(k)+\frac{1}{16}x_3^C(k)+\frac{17}{12}\right)\\ x_3(k) &= f\left(\frac{1}{16}x_1^C(k)+\frac{1}{32}x_2^C(k)+\frac{1}{8}x_3^C(k)+\frac{1}{2}\right)\end{aligned}\right\}$$
$$k \equiv 0 \pmod 3$$

It suffices to find the initial values $x(0), x^C(0)$, so that $x(3) = x(0), x^C(3) = x^C(0)$. We have found that, approximately:
$x_1^C(0) = 1.8403,$
$x_2^C(0) = 1.614,\ x_3^C(0) = 1.3689,$
$x_1(0) = 0.9705, x_2(0) = 0.89425,$
$x_3(0) = 0.77054$
(see Table 2 for the first 4 values of the 3-periodic solution of the above system).

Table 2. A 3-periodic solution in Case b).

	$k=0$	$k=1$	$k=2$	$k=3$
$x_1^C(k)$	1.8403	1.89065	1.813606	1.8400157
$x_2^C(k)$	1.614	1.70125	1.5821907	1.6145067
$x_3^C(k)$	1.3689	1.45499	1.3448186	1.3689537
$x_1(k)$	0.9705	0.868281	0.9332127	0.9705381
$x_2(k)$	0.89425	0.7315657	0.8234114	0.8943396
$x_3(k)$	0.77054	0.6173236	0.6965444	0.7705772

The initial values in Cases a) and b) have been found using MATLAB, after numerous experiments with different sets of possible initial values.

4. Conclusions

We found a periodic solution of a modified Elman neural network satisfying the sufficient conditions. In the future, we plan to consider similar problems for Output Hidden Feedback Elman neural networks. The model considered can be applied to the services quality of experience prediction.

References

[1]. J. L. Elman, Finding structure in time, *Cognitive Science*, Vol. 14, Issue 2, 1990, pp. 179-211.

[2]. G. Ren, Y. Cao, S. Wen, T. Huang, Z. Zeng, A modified Elman neural network with a new learning rate scheme, *Neurocomputing*, Vol. 286, 2018, pp. 11-18.

[3]. L. Xu, Y. Zhang, Quality prediction model based on novel Elman neural network ensemble, *Complexity*, Vol. 2019, 2019, 9852134.

[4]. T. Xia, P. Zhuo, L. Xiao, S. Du, D. Wang, L. Xi, Multi-stage fault diagnosis framework for rolling bearing based on OHF Elman Adaboost-Bagging algorithm, *Neurocomputing*, accepted.

[5]. X. Gao, D. You, S. Katayama, Seam tracking monitoring based on adaptive Kalman filter embedded Elman neural network during high-power fiber laser welding, *IEEE Transactions on Industrial Electronics*, Vol. 59, Issue 11, 2012, pp. 4315-4325.

[6]. M. Trifonova, X. Li-Shtereva, Analytical mathematical models for determining the probable sliding surface of the working slope, in *Proceedings of the International Symposium Mining and Geology Today*, Belgrade, 2017, pp. 179-182.

[7]. S. Ding, Y. Zhang, J. Chen, W. Jia, Research on using genetic algorithms to optimize Elman neural networks, *Neural Computing and Applications*, Vol. 23, 2013, pp. 293-297.

[8]. W. Z. Sun, J. S. Wang, Elman neural network soft sensor model of conversion velocity in polymerization process optimized by chaos whale optimization algorithm, *IEEE Access*, Vol. 5, 2017, pp. 13062-13076.

[9]. L. Yang, F. Wang, J. Zhang, W. Ren, Remaining useful life prediction of ultrasonic motor based on Elman neural network with improved particle swarm optimization, *Measurement*, Vol. 143, 2019, pp. 27-38.

[10]. V. Covachev, Z. Covacheva, Existence of periodic solutions for a modified Elman neural network, in *Proceedings of the AIP Conference on New Trends of Differential Equations in Sciences (NTADES'20)*, Sts. Constantine and Helena, Bulgaria, 1-4 Sep. 2020, accepted.

[11]. X. H. Shi, Y. C. Liang, H. P. Lee, W. Z. Lin, X. Xu, S. P. Lim, Improved Elman networks and applications for controlling ultrasonic motors, *Applied Artificial Intelligence*, Vol. 18, Issue 7, 2004, pp. 603-629.

[12]. A. Berman, R. J. Plemmons, Nonnegative Matrices in Mathematical Sciences, *Academic Press*, New York, 1979.

[13]. M. Fiedler, Special Matrices and Their Applications in Numerical Mathematics, *Martinus Nijhoff*, Dordrecht, 1986.

[14]. R. E. Gaines, J. L. Mawhin, Coincidence Degree and Nonlinear Differential Equations, *Springer-Verlag*, Berlin, 1977.

(007)

Predictive Control with Energy Efficiency Enabled by Real-time Machine Learning

G. Y. Luo [1, 2], Y. Q. Luo [3] and H. G. Gan [1]
[1] School of Physics and Materials Science, Guangzhou University, Guangzhou 510006, China
[2] SHJ Medical Gas Specialists, Asheridge Road, Chesham HP5 2QA, UK
[3] Department of Mathematics, London School of Economics and Political Science, London WC2A 2AE, UK
E-mail: gaoyong.luo@yahoo.com

Summary: Through the analytics of real-time data collected from internet of things (IoT) sensor networks, machine learning system has enabled confident decision making and made it possible to optimise the controls and maintenance of the physical assets, manufacturing systems and processes of prediction and prevention, and further to achieve maximum power efficiency and productivity. In many industrial applications, it is critical to use deep neural networks that make predictions both fast and accurate. Indeed, deep learning is a foundational technology for predictive control and maintenance that can be applied to coupled multiple-input multiple-output (MIMO) systems of industrial processes. However, due to the strong nonlinearity and its nearly instantaneous response to disturbances, it is still very challenging to achieve predictive control with optimised performance in such a complex MIMO system. In this paper, to raise the power and capabilities of machine learning, we propose a novel machine learning system designed by building IoT networks to remotely collect data and developing deep wavelet neural networks (DWNN) with Gaussian (Mexican hat) wavelet derived as activation functions to improve nonlinear fitting and convergence speed, and to process the real-time data for predictive control with energy efficiency. Experimental results demonstrate that optimal performance of MIMO system operations can be achieved by the trained DWNN over the IoT communication networks, where the sensing capabilities and the computational power are provided by the designed controller, transmitter and cloud server to track everything that is relevant to operations, such that by deep learning with real-time data analytics we can have a knowledge base from which to correct errors, improve control strategy and maximise efficiency.

Keywords: Real-time machine learning, Complex MIMO systems, Deep wavelet neural network, Activation function, Data analytics, Predictive control, Energy efficiency.

1. Introduction

With the advances in artificial intelligence (AI) and machine learning (ML), real-time data from various sensors and intelligent systems over the internet of things (IoT) networks can be remotely collected and processed to monitor machine conditions, detect faults, predict and optimise control processes, which offer ways to optimise the operations and maintenance of the physical assets, manufacturing systems and processes of prediction and prevention, by enabling the technological innovation that the industry needs. Equipped with many sensors over the IoT networks and enabled by deep learning to detect relevant status information of the actuators and to monitor the health status of specific components [1], this new data-driven model is improving and dramatically changing the way how predictive control system can be designed and that energy efficiency can be achieved through minimising power consumption. The resulting cost savings and competitive advantages are essential as the evolution and convergence of many new technologies – mechatronic systems, controllers, edge and cloud computing, big data, machine learning and the IoT, are adopted and providing the basis for increasing the self-awareness of the machine, allowing it to optimise its own performance for the given duty cycles, diagnose and compensate for non-catastrophic faults, and coordinate operations with other machines with minimal input from the operator. From the perspective of production and service management, establishing the intelligent and communicative systems can enable the processes to deal with the data flow from intelligent and distributed system interaction and to promote autonomous interoperability, agility, flexibility, decision-making, and efficiency [2].

Through enabling controllers to perform specific tasks intelligently, machine learning systems can carry out complex processes by learning from data, rather than following the pre-programmed rules, allowing a compelling and robust control architecture with optimised performance to be built on [3]. Automatic learning also vastly improves product quality by introducing predictive maintenance systems into production processes, replacing visual inspections that can execute quality controls more accurately and efficiently. Indeed, deep learning is regarded as a foundational technology for complex applications such as predictive control and maintenance, and can be applied to coupled multiple-input multiple-output (MIMO) systems of industrial processes. However, due to the strong nonlinearity and its nearly instantaneous response to disturbances, the predictive control with optimised performance in such a complex MIMO system is difficult to achieve [4]. It is required to build up appropriate research environments and obtain datasets that can be used to explore and raise the power and capabilities of machine learning, where commonly used deep neural networks require improvements in nonlinear fitting and convergence

speed. In this paper, we propose a novel machine learning system designed by building IoT network to remotely collect data and developing deep wavelet neural networks using nonlinear Gaussian (Mexican hat) wavelet with sparsity as activation functions to process the real-time data for predictive control with energy efficiency.

2. Machine Learning System

Real-time machine learning system is designed as shown in Fig. 1, where by designing printed circuit boards (PCBs) as machine controllers, signal collectors and transmitters, the IoT communication networks could be built for real-time monitoring, control and optimising by machine learning, so that the predictive control and maintenance for highly efficient operation of industrial systems can be achieved. This means the current systems can be upgraded to suit next generation needs of actioning changes in a process to optimise the efficient use of industrial systems, thus creating opportunities for industry 4.0 to be implemented with modernised products, systems and services, and also to improve productivity, precision, flexibility, and profitability. By real-time machine learning, the designed controller is able to predict and optimise control parameters to minimise energy consumption, while by real-time processing of measurement data provided by dedicated sensors installed in the machine, the system can enable autonomous decision making based on the online diagnosis of the correct machine with right condition, leading to increased machine reliability towards zero defects, together with higher productivity and efficiency. Predictive maintenance based on condition monitoring by measuring such as vibration, acoustic, motor current etc., can then be planned and scheduled, such that we can have a knowledge base from which to correct errors and improve efficiency.

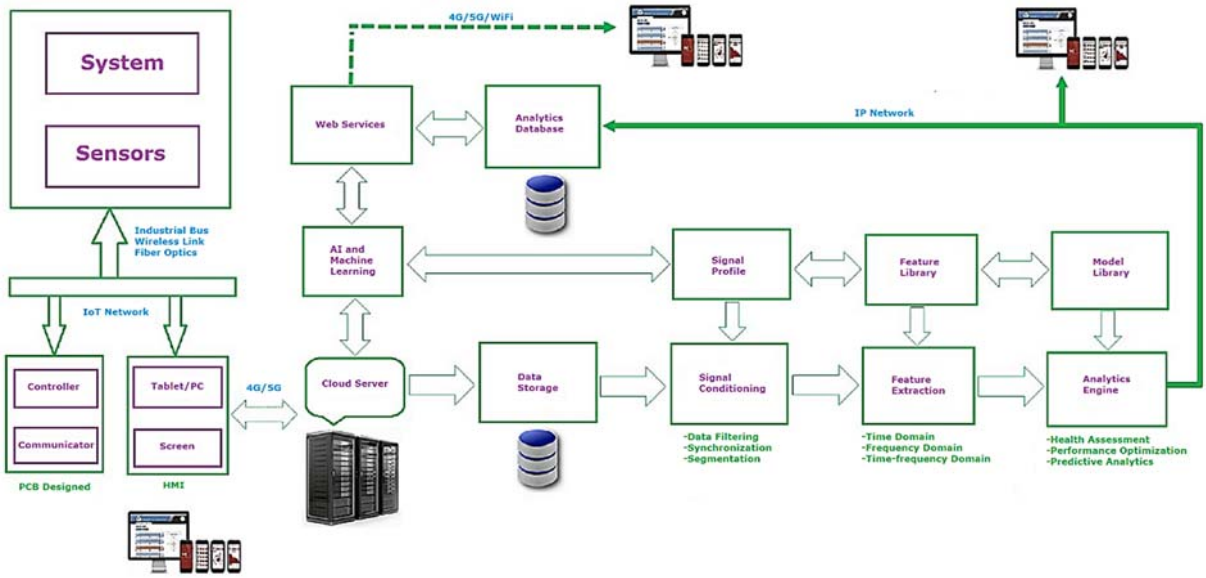


Fig. 1. Real-time machine learning system.

For MIMO systems of industrial processes, they can be modelled by the form of regression problem aimed at predicting:

$$Y = f(X) + \epsilon, \tag{1}$$

where $f(X)$ is an unknown function with inputs X and outputs Y, and ϵ is typically white additive noise processes represented by a random matrix. Prediction can be achieved by the nonlinearly mapping or fitting that models the observed past data, and we aim to minimise the mean squared error (MSE) between the actual inputs and predicted outputs at the training process.

For one input X_i and one output Y_m, we have

$$Y_m = f(X_i) + \epsilon_m d$$

If $X_i = (x_t, x_{t-1}, x_{t-2}, \cdots, x_{t-j})$ represents the past data, and $Y_m = (y_{t+1}, y_{t+2}, \cdots, y_{t+k})$ denotes the prediction with k steps, the time-series forecasting would be:

$$(y_{t+1}, y_{t+2}, \cdots, y_{t+k}) = \\ = f(x_t, x_{t-1}, x_{t-2}, \cdots, x_{t-j}) + \epsilon_m, \tag{2}$$

where $i, j, k, m \in$ Z (Z is the set of all integers). The prediction is a complex nonlinear mapping where deep wavelet neural networks (DWNN) with nonlinear fitting can be used for predictive control and maintenance.

3. Deep Wavelet Neural Networks

To better map the inputs to outputs with highly nonlinear relationship, we construct deep wavelet neural networks (DWNN) as shown in Fig. 2.

The prediction output Y_m can be expressed by deep prediction rule:

$$\begin{aligned} Y^{(1)} &= f^{(1)}\left(W^{(1)}X_i + B^{(1)}\right), \\ Y^{(2)} &= f^{(2)}\left(W^{(2)}Y^{(1)} + B^{(2)}\right), \\ &\vdots \\ Y^{(L)} &= f^{(L)}\left(W^{(L)}Y^{(L-1)} + B^{(L)}\right), \\ Y_m &= W^{(L+1)}Y^{(L)} + B^{(L+1)} \end{aligned} \tag{3}$$

where f is a wavelet function that is used as activation function to construct the DWNN, W is the weight matrix and B denotes bias. Rewrite equation (2) by the analysis of variance (ANOVA) expansion [5]:

$$\begin{aligned} y(t) &= f\left(x_1(t), x_2(t), \cdots, x_n(t)\right) + e(t) = \\ &= f_0 + \sum_{i=1}^{n} f_i\left(x_i(t)\right) + \\ &+ \sum_{1\le i<j\le n} f_{ij}\left(x_i(t), x_j(t)\right) + \\ &+ \sum_{1\le i<j<k\le n} f_{ijk}\left(x_i(t), x_j(t), x_k(t)\right) + \cdots + \\ &+ \sum_{1\le i_1<\cdots<i_m\le n} f_{i_1 i_2 \cdots i_m}\left(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_m}(t)\right) + \\ &+ \cdots + f_{12\cdots n}\left(x_1(t), x_2(t), \cdots, x_n(t)\right) + e(t) \end{aligned} \tag{4}$$

where $i_1, i_2, \cdots, i_m, m, n \in \mathrm{Z}$.

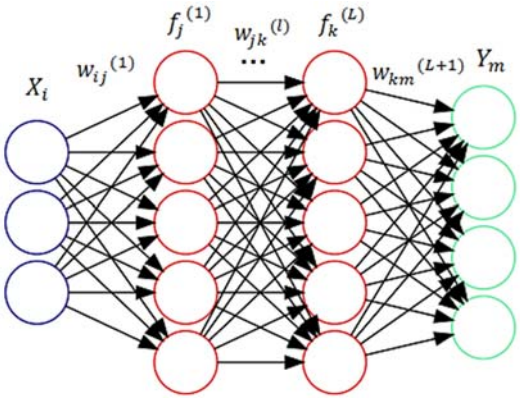


Fig. 2. A deep wavelet neural network (DWNN).

For any function $f \in L^2(\boldsymbol{R})$, the form of wavelet decomposition is

$$\begin{aligned} f(x) &= \\ &= \sum_j \sum_k c_{j,k}\psi_{j,k}(x) = \sum_{m=1}^{M} c_m \psi_m(x) \end{aligned} \tag{5}$$

For an n-dimensional wavelet

$$\psi^{[n]} = \psi^{[n]}(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} \psi(x_i)$$

If we choose Gaussian wavelet functions, we have

$$\begin{aligned} \psi^{[n]} &= \psi^{[n]}(x_1, x_2, \cdots, x_n) = \\ &= x_1 x_2 \cdots x_n e^{-\frac{\|x\|^2}{2}}, \end{aligned} \tag{6}$$

where $\|x\|^2 = \sum_{i=1}^{n} x_i^2$. In this study, we use Mexican hat wavelet as activation functions. It is the negative normalised second derivative of a Gaussian function and has admissibility condition and a symmetric structure.

$$\psi_{mexh}(t) = \frac{2}{\sqrt{3}\pi^{\frac{1}{4}}}(1 - t^2)e^{-\frac{t^2}{2}} \tag{7}$$

Its derivative is also a Gaussian function:

$$\psi'_{mexh}(t) = \frac{2t}{\sqrt{3}\pi^{\frac{1}{4}}}(t^2 - 3)e^{-\frac{t^2}{2}}$$

Thus the expansion component $f_{i_1 i_2 \cdots i_m}(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_m}(t))$ can be expressed by DWNN model:

$$\begin{aligned} &f_{i_1 i_2 \cdots i_m}\left(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_m}(t)\right) = \\ &= \sum_{j=j_m}^{J_m} \sum_{k_1 \in K_j} \cdots \\ &\sum_{k_m \in K_j} c_{j;k_1,\cdots,k_m} \psi_{j;k_1,\cdots,k_m}^{[m]}(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_m}(t)) \end{aligned} \tag{8}$$

where $k_1, k_2, \cdots, k_m, j_m, J_m, K_j \in \mathrm{Z}$.

Compared with commonly used deep neural networks, using nonlinear Gaussian (Mexican hat) wavelet with sparsity as activation functions in the DWNN model can improve nonlinear fitting and convergence speed with adaptive learning rate, so that the power and capabilities of machine learning can be raised.

4. Experiments

To evaluate the proposed DWNN for predictive control and maintenance, unmanned medical air plant was set up as shown in Fig. 3, where it consists of 4 compressors and 2 dryers to compress and purify air for hospital users.

Based on the choice of fixed speed compressor, with condition monitoring for medical compressed air plant and real-time machine learning for predictive control, the plant can achieve similar performance as variable speed compressor in terms of energy saving but with lower costs. The dynamic flow control of the plant is to keep the pressure and quality of output air at the level required by medical usage, which can be assured by measuring the receiver pressure (defined as P) and dewpoint temperature (defined as T). As for predictive control, parameters P and T are used to determine the control logic in order to balance the air flow between input and output, where dryers generating process at the time of purging is included as compressed air lost (purge lost). Now define the input air flow rate as v_i (generated by the duty compressor(s), output as v_o (used by surgical needs), and purging air flow rate at dryer 1 as v_{d1}, and dryer 2 as v_{d2}, the compressed air pressure equalising function can then be expressed as

$$\begin{aligned} &\int v_i(t)dt \cdot e^{-(t-\tau)} \triangleq \int v_o(t)dt + \\ &\int v_{d1}(t)dt + \int v_{d2}(t)dt \end{aligned}$$

where τ denotes time delay when input and output air pressure are balanced, and can be identified by machine learning at the commissioning stage. The parameter v_{d1} and v_{d2} can also be estimated at this stage by performing tests with the system modelled

and trained by DWNN algorithm. To ensure all compressors and dryers provide optimum performance, predictive control strategy is employed with system parameters P and T measured and predicted. After training the DWNN by minimising the error between the actual outputs and desired outputs to determine the network parameters, it is used online to continuously predict the control parameters based on the past measured data to optimise the control strategy. An example of time series pressure (parameter P) measurement and prediction by the DWNN is shown in Fig. 4 which can be used for predictive control. The DWNN training and processing is performed at the cloud server site with optimised control parameters adjusted and sent back to the edge, where each PCB controller of compressor is collecting signals of collective faults and analogue input of pressure and motor current transducers, while one of the PCB controllers of dryer is collecting switch signals of pressure low and dewpoint temperature high, and analogue input of receiver pressure (P) and dewpoint temperature (T) transducers. A central PCB controller connected is taking the role of master and computing for machine sequence control with general pressure control between 9.0 and 10.0 bar (lower pressure control with less purge lost) as follows (Table 1).

Table 1. Machines sequence control.

Compressor on duty	Dryer on duty	Time of operating
Loading	Drying/Operating	4 minutes
Idling	Equalising	1 minute
Loading	Regenerating/Purging	4 minutes
Idling	Equalising	1 minute

By the real-time machine learning system built, predictive control and condition monitoring based maintenance with real-time data measured from sensors of motor current, vibration and acoustic, can then be performed by mapping the multiple inputs to outputs using the DWNN framework for high quality and consistency of air supply with energy efficiency, keeping costs to a minimum. From the tests performed, it is noted that up to 30 % of energy costs can be saved through real-time data analytics by the DWNN developed.

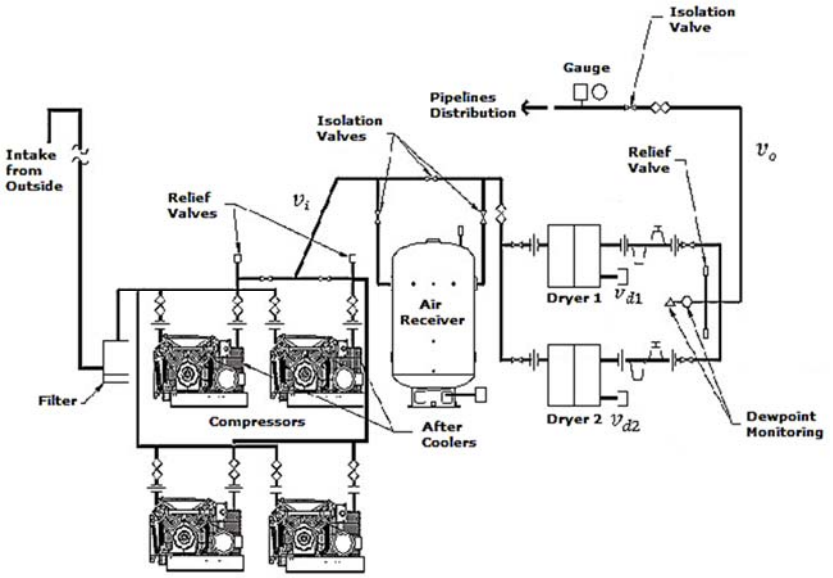


Fig. 3. Unmanned medical air plant.

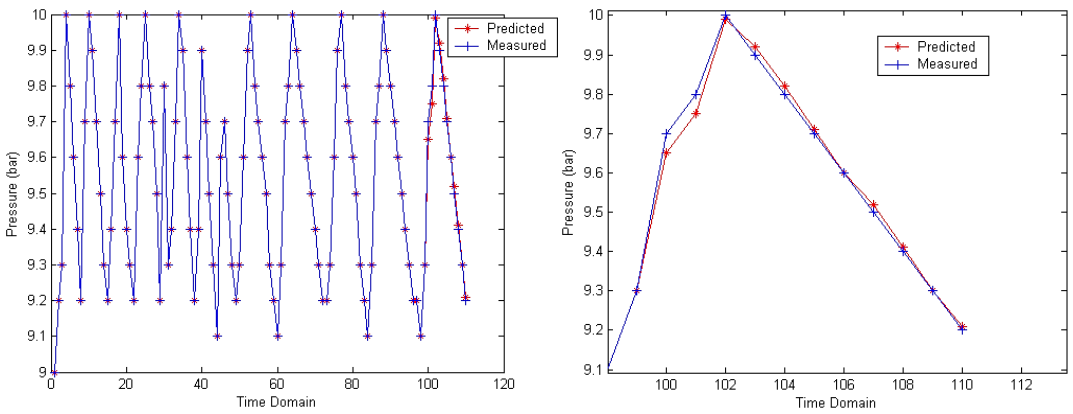


Fig. 4. Time series pressure (parameter P) measurement and prediction.

5. Conclusions

Equipped with many sensors over IoT networks and enabled by deep learning to detect relevant status information of the actuators and to monitor the health status of specific components, the new data-driven model proposed is improving and dramatically changing the way that predictive control system can be designed and energy efficiency can be achieved through minimising power consumption. Through the analytics of real-time data collected from internet of things (IoT) sensor networks, machine learning system enables confident decision making and predictive control and maintenance, and can be applied to coupled multiple-input multiple-output (MIMO) systems of industrial processes. However, due to the strong nonlinearity and its nearly instantaneous response to disturbances, predictive control with optimised performance in complex MIMO systems is difficult to achieve. It is required to build up appropriate research environments and obtain datasets that can be used to explore and raise the power and capabilities of machine learning, where the commonly used deep neural networks require improvements in nonlinear fitting and convergence speed. In this paper, we propose a novel machine learning system designed by building IoT network to remotely collect data and developing deep wavelet neural networks (DWNN) using nonlinear Gaussian (Mexican hat) wavelet with sparsity as activation functions to process the real-time data for predictive control with energy efficiency.

Experimental results have demonstrated that optimum performance of MIMO system operations can be achieved by the trained DWNN model over the IoT communication networks, where the integrated high technologies such as IoT, machine learning, cloud computing enable systems to make the best decisions, improve control strategy and maximise efficiency.

References

[1]. K. Patan, Neural network-based model predictive control: Fault tolerance and stability, *IEEE Transactions on Control Systems Technology*, Vol. 23, Issue 3, 2015, pp. 1147-1155.

[2]. V. Alcácer, V. Cruz-Machado, Scanning the Industry 4.0: A literature review on technologies for manufacturing Systems, *Engineering Science and Technology, An International Journal*, Vol. 22, 2019, pp. 899-919.

[3]. A. Draeger, S. Engell, H. Ranke, Model predictive control using neural networks, *IEEE Control Systems Magazine*, Vol. 15, Issue 5, 1995, pp. 61-66.

[4]. T. Baumeister, S. L. Brunton, J. N. Kutz, Deep learning and model predictive control for self-tuning mode-locked lasers, *Journal of the Optical Society of America B*, Vol. 35, Issue 3, 2018, pp. 617-626.

[5]. S. A. Billingsand, H. L. Wei, A new class of wavelet networks for nonlinear system identification, *IEEE Transactions on Neural Networks*, Vol. 16, Issue 4, 2005, pp. 862-874.

(008)

On Brain and Cognitive Intelligence Based Control in Robotics

B. Wei
Algoma University, 1520 Queen St E, Sault Ste Marie, Canada
Tel.: + 1 705-949-2301, ext. 4270
E-mail: bin.wei@algomau.ca

Summary: In control arena, one of the most important issues is to effectively control the robot motion and force and in the meantime, the robots must safely respond to contact forces while interacting with people. Another issue is that how to handle uncertainties that are occurred during the process of interacting with environment. Human beings can control ourselves' motions (e.g. control arms to move, legs to move, whole body to move) perfectly. This partly attributes to the fact that human control system is based on millions of neurons that are receiving and sending signals to each other, and partly attributes to the fact that this control is formed by our initial learning. This paper reviews the most recent development on the brain and cognitive intelligence based control used in robotics. Some future research recommendations are proposed. It is concluded that the final goal is how to combine neuroscience, AI, and robotics so that we can make robots have more human-like performances.

Keywords: Robotics, Adaptive control, Cognitive intelligence, Learning control, Human-robot interaction.

1. Introduction

In 2000, the humanoid robot ASIMO is developed by Honda research group, and the robot can walk and climb stairs etc., but the robot cannot have physical interaction with humans [1, 2]. Recently, the Boston Dynamics research group together with Google developed a humanoid robot Atlas, surprisingly it is not heavily based on artificial intelligence but motor control and learning through the state space model, and it is by far the most agile robot on the market [3, 4]. However, the Atlas robot was not designed to have physical interaction with human beings, so it cannot really have physical interaction with human beings. Similarly, the recent developed humanoid robot Sophia that is largely based on artificial intelligence developed by Hanson Robotics research group can have verbal interactions with human beings, but not too much physical human robot interaction.

Before proceeding further, it is noted that biologically inspired robots have been developed in the past decades [5-7], in which researchers are trying to simulate and copy biology organisms to make robots more intelligent and accomplish complicated tasks. What about simulating and copying human behaviors (human nervous system) instead of animals behaviors? Since the most reliable and intelligent control system ever encountered is the human internal control system.

Human beings can control ourselves' motions (e.g. control arms to move, legs to move, whole body to move) perfectly. This partly attributes to the fact that human control system is based on millions of neurons that are receiving and sending signals to each other [8-10], and partly attributes to the fact that this control is formed by our initial learning [11, 12]. It is noticed that when we were babies, we learned how to control the motions over and over again, and gradually formed the perfect control of movement.

Some of the traditional control systems (e.g. PD control) mimic the spring and damper system, as briefly illustrated in Fig. 1. Traditional control systems (used in robots) that are mainly geared to the industrial manufacturing purpose are no longer effective in human-robot interaction, e.g. one of the applications is eldercare. As serving an elderly people requires more complicated motions and safety issues need to be taken into consideration as well while the robot is interacting with elderly [13, 14]. The robot requires more than just a position control, motion control, force control, or a trajectory planning [15-17]. It also involves with combination of above and unexpected motion control [18, 19]. For example, in the process of assisting an elderly to walk around, if an elderly accidently falls, the robot needs to handle this unexpected motion and help to pick up the elderly. A learning control system inspired by human brain and cognitive intelligence system seems to be a good candidate to address the above problem, i.e. how to design a human-nervous inspired learning control algorithm for assistive robots to help elderly people in their daily lives, for example, help to pour a cup of water and deliver it to the elderly, help to assist elderly to walk around, and help to clean the house. All these tasks need a robot and an advanced control algorithm. One of the most important issues is to effectively control the robot motion and force and in the meantime, the robots must safely respond to contact forces while serving an elderly people [20, 21]. Another issue is how to handle uncertainties that are occurred during the process of, for example, serving an elderly people [22, 23]. The development of physical assistive robots is not a main issue here, the main issue is to develop an advanced control algorithm to control

the robots to execute an unexpected tasks in the process of assisting elderly people [24-26].

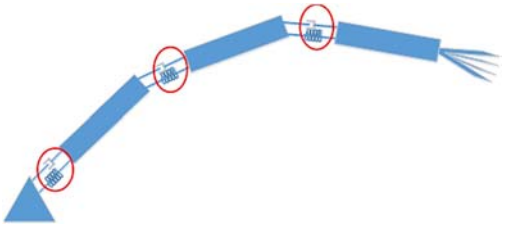


Fig. 1. PD control analogy.

Here, the most recent development of brain and cognitive intelligence based control are reviewed and the human-nervous-systems-based learning control algorithm for robots is discussed in order to further summarize and improve the methodologies in this field.

2. Cognitive Based Control in Robotics

In [27], the author studied how human beings control arms' movements by looking at the patterns of electrical activity in neurons when a person moves his or her arm through combining artificial intelligence and neuroscience, and then apply such skills to robots. The studies show that how artificial intelligence and neuroscience can help each other in order to advancing both fields.

Very similarly, artificial intelligence, computer programs, and brain organization are compared and investigated together in [28], and it came to the same conclusion that AI and brain organization are helpful to each other. The output of brain organization might be useful in better modelling and simulating of AI. In [29], a deep neural network is identified and trained by computational approaches in order to study the patterns of electrical activity of a monkey neurons in the process of recognizing an object. The deep neural network was modelled based on the human brain hierarchical architecture, and it is acknowledged that generally neural networks perform supervised learning. In [30], a "brain-like" control technique is used in the visual servo control of a redundant 7-DOF robotic manipulator for the tele-robotic operations application to remove the image-feature extraction and tracking requirements. However, the technique requires installation of advanced sensing systems and reconfigurable base. Furthermore, the force control is not considered and also safety issue is one of the major concerns when applying above approach. Precisely speaking, the above brain-like control technique can actually be considered as non-vector space control.

In [31, 32], the intersection of cognitive science, computational neuroscience, and artificial intelligence are reviewed. It shows that by using Bayesian cognitive models, one can engage complex cognition, such as the way our minds model the physical world by optimally combining prior knowledge about the world with sensory evidence. However, the study did not specify how we can apply such approach to robots. In [33], the authors discussed how to build and make robots that learn like people. The article presented and illustrated two major challenges in the process of robot learning: image and character recognitions, and learning how to play a video game. The authors argued that the combination of neural network modeling and learning-based artificial intelligence could result in much better result and more human-like learning abilities than that of single artificial intelligence system. Furthermore, it also concludes that deep neural networks learning together with psychological ingredients could lead to a much higher level of cognitive control. The challenging problem and of course the next step of work seem to be that how to build the bridge between deep neural networks learning and psychological ingredients.

There are also numerous of recent studies focusing on the deep learning and neuroscience [35-40], but the studies did not specify how exactly to connect deep learning and neuroscience to robot controls.

In [41], a recurrent emotional cerebellar model articulation controller neural network is presented for vision-based mobile robots to handle complex nonlinearities in the dynamic equations so as to achieve accurate positioning tasks. The developed control neural network combines a recurrent loop with an emotional learning system and therefore generate a cerebellar model articulation control system, which is acted as the main piece of the control system module. The developed control system contains three parts: a sliding surface, the recurrent emotional cerebellar model articulation controller, and a compensator controller, and the control system is designed based on the Lyapunov stability theory and therefore, the stability of the overall controller is guaranteed. However, the designed control network has certain limitations where the control network can only be used in target tracking due to the fact that lacking of self-organization mechanism in the overall control system. In [42], the authors proposed an approach by training machine learning regression algorithms to anticipate slippage related to individual wheels in mobile robots that are used in off-road conditions. The novelty about that study is that the machine learning regression is employed instead of machine learning classification to detect slippage and related uncertainties. The interesting part is that the slippage is considered as a random variable. The MIT single-wheel equipment is used as a demonstration, and the result shows that the Gaussian process regression leads to a trade-off among computation time and precision. Using deep learning approaches to model slippage as a discrete variable will be interesting to look at in the future.

In [43], a framework for investigating the nature of consciousness in cognitive robots is developed, and the issue of how to create a humanoid robot that learns to conduct works based on imitation learning is discussed. It is pointed out that the above humanoid cognitive robot can be achieved by copying and following human-provided demonstrations (i.e. the

so-called learning from demonstrations) rather than pre-programming a robot. However, the potential issue is that the robot is very limited to generalize to a completely new set of circumstances where the robot needs to use different actions to execute the same intention, because the robotic system does not understand the demonstrator's intentions. In [44], the development of neuroscience and psychology-inspired human – robot interaction is briefly reviewed, and some challenging issues are brought up. The study reviewed the above topic from the following three different perspectives: technical solutions for human-robot interaction, the development of social behaviour in robots, and human social perception and social behaviour. Nevertheless, it was concluded that we need to integrate different research areas to make significant progress in controls of human-robot interaction.

In [45], a framework for coordinating stimulus sport interaction scenarios with robots is developed based on motivational instruction patterns. It is shown that by employing a structured way of assessing multiple interaction configurations and robot platforms, one can obtain a deep understanding in psychological concepts that shape human-robot interaction. However, the study is still preliminary in terms of scenarios presented.

In the beginning of the century, people are looking at the adaptive control to handle the situation where parameters in the dynamic equation change with respect to time. Adaptive control still replies on models and it is a type of model-based control. Later, many research groups investigated the model-free control, particularly the deep reinforcement learning in robotics [46-48]. The challenging problem we are facing is that how to make robots deal with uncertainties and make corresponding decisions, and how to make robots truly achieve human-like intelligence.

In [49], the authors investigated the human–robot cooperative tasks by looking at the needed individual, geometric reasoning and situation assessment based on affordance analysis and perspective-taking. The context is that robots and humans are put in a common space, and humans and robots exchange information mainly via verbal communication and social observation, and the robot is then anticipated to carry off interactive object manipulation by considering the human's intentions, skills and beliefs. The study presented is till preliminary. It does not consider cases where different beliefs might appear between the robot and human. Moreover, the study is only limited at the spatial and temporal situations, where a context selection consists of retrieving one monolithic set of beliefs rather than single one belief. In another interesting and similar study [49], a coordinated speech-based humanoid robot is developed that can follow human commands and produce poems, receive audiences feedback and thereafter display corresponding emotional responses, and produce body language movements accordingly. The dialogue capabilities of the robot yet needs to be further improved in order to allow an active learning for extracting and organizing knowledge from human.

In programming arena, for example in Java, we have the concept of modularized programming (Robert 2008), where one lengthy program is reduced to one main method and several blocks of sub-methods. Each of these sub-methods is one module, and each module does its own work and returns the result to the main method which utilizes the returned result. A Java example is shown here in Fig. 2 for the purpose of an illustration. Furthermore, it is argued that our human body does not only have one controller, but rather multiple controllers that are working together with the central nervous system, which is our brain and spinal cord. When humans control bodies, not all controllers are working at the same time. For example, when human humans grab an object, the only control system working might be the arm controller and other related controller, the controller that controls our facial expression might not be working. Or as another example, when human beings lay on the bed and relaxing, we still have all the controllers but most of the controllers are not performing to control any movement. The idea of the multiple controllers working independent and together with the central control system might be one of the future research directions we can take to handle robot control in order for robots to accomplish more complex and delicate works, as conceptually demonstrated in Fig. 3. We can build a model to represent such hierarchy concept and apply such control to robots.

3. Discussions

The general approach as a future study that we are going to take for designing cognitive based control for robotics or human-nervous-system-based learning control algorithm is to build a mathematical model to resemble and represent human nervous system based on the structure and function of the nervous system, and then apply such model to robots.

The following states how we are going to achieve human-nervous-system-based learning control algorithm in details: in the designing of such human-nervous-system-based learning control algorithm, the initial idea is to make the human-nervous-system-based learning control system have two major sections. The first section contains the learning control algorithm. The learning control algorithm can be designed based on the artificial intelligence by simulating the human learning approach (learning over and over again) and modelling a deep neural nets model. After that, maybe we can add something that is very similar to the structure of neurons that are receiving and sending signals to each other to the above learning control algorithm, and this forms the second section, which will be one of our future works. The purpose of the first section (the learning control algorithm) is to store enough "correct move" and "wrong move" information. The purpose of the second section (the neurons alike structure) is to

utilize the first section's information to communicate with the robot. Fig. 4 briefly illustrates the overall plan for designing a human-nervous-system-based learning control algorithm used, for example, in a socially assistive robot. The key problem that needs to be solved is the human-nervous inspired learning control algorithm instead of building a robot.

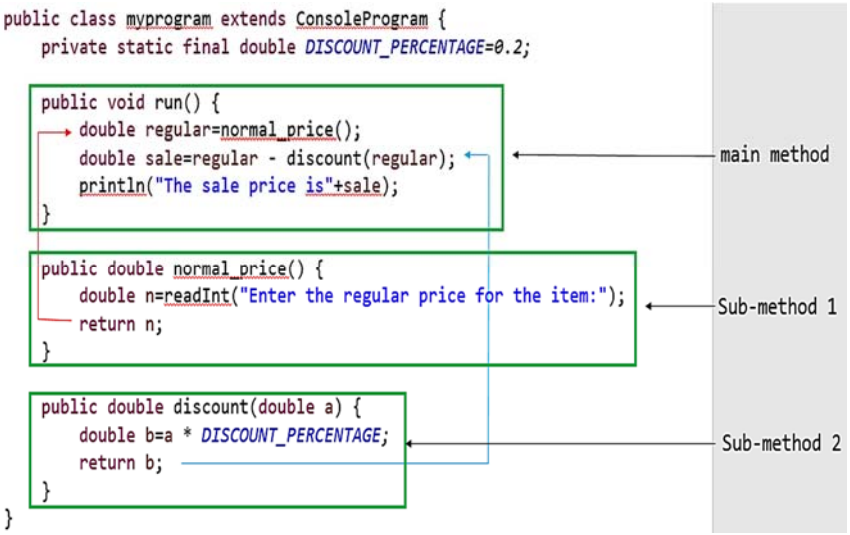


Fig. 2. A Java program example – modularized programming concept.

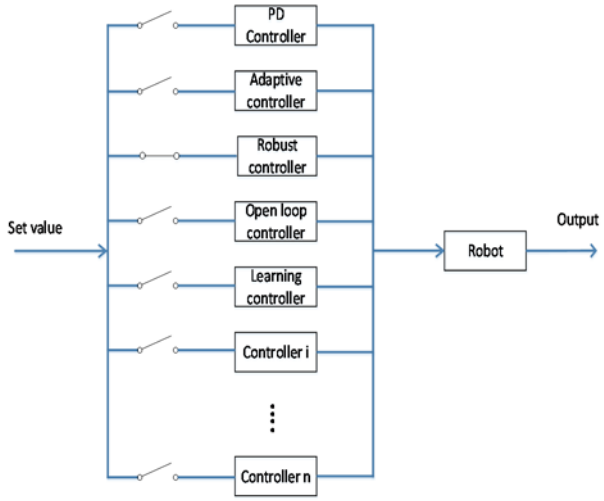


Fig. 3. Multiple controllers work together and independently to control a robot.

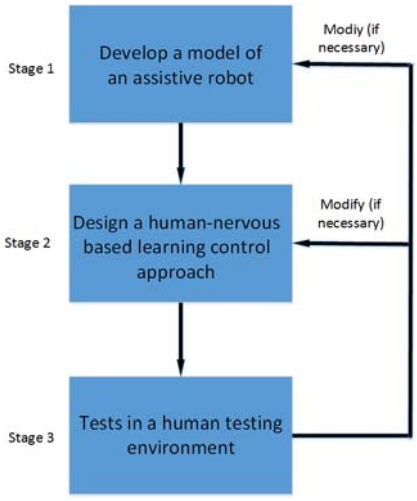


Fig. 4. Example of human-nervous inspired learning control algorithm design overall plan.

4. Potential Applications

Robotic systems have been used in many industrial areas, but its development in the human-robot interaction, for example, socially-assistive robots in eldercare, is still in the infant stage [51]. It is estimated that by 2030, there will be over 1 billion elderly people who are over 60 years old around the world. Young people will not be able to stay at home and serve their older parents often as they have to go to work. Thus the use of assistive robots in eldercare will become necessary in the near future. There are few robots being designed in the last decade where the robots can do some basic tasks for elderly, but mainly on the communication and entertainment level [52-54], not actually having physical interactions with elderly. The significance and potential impact of the cognitive based control research are summarized as follows:

(1) By 2030, it is estimated that 20 % of world's population will be over 60 years old. The development of socially-assistive robots will greatly assist elderly people in their everyday lives;

(2) The development of human-nervous inspired learning control algorithm will further push the development of the automatic control theory to a new level and introduce a new concept of control, which can potentially be used also in advanced manufacturing. Manufacturing plays a crucial role in

today's economy responsible for about 16 % of global gross domestic products. In current manufacturing industries (e.g. in manufacturing assembly lines), robots are working independently and not working with human operators. In order to improve the productivity and adapt to the rapid changes in production, current assembly robots are very limited in satisfying those needs. Whereas robots working with human operators is an effective approach to meet the needs of rapid changes in production.

5. Conclusions

The author here reviewed the most recent development of the brain and cognitive intelligence based control used in robotics in order to further summarize and improve the methodologies in this field. A human-nervous-systems-based learning control algorithm design framework for assistive robots is briefly illustrated and discussed. Some future research recommendations are proposed. It seems the final goal is that how to combine neuroscience, AI, and robotics so that we can make robots have more human-like performances.

References

[1]. J. Kuffner, K. Nishiwaki, S. Kagami, M. Inaba, H. Inoue, Footstep planning among obstacles for biped robots, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'01)*, 2001, pp. 500-505.

[2]. Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, K. Fujimura, The intelligent ASIMO: system overview and integration, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'02)*, Vol. 3, 2002, pp. 2478-2483.

[3]. E. Ackerman, E. Guizo, The Next Generation of Boston Dynamics' ATLAS Robot Is Quiet, Robust, and Tether Free, http://spectrum.ieee.org/automaton/robotics/humanoids/next-generation-of-boston-dynamicsatlas-robot

[4]. S. Feng, E. Whitman, X. Xinjilefu, C. G. Atkeson, Optimization based full body control for the atlas robot, in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, Spain, 2014, pp. 120-127.

[5]. H. Faraji, R. Hatton, Aiming and vaulting: Spider-inspired leaping for jumping robots, in *Proceedings of the IEEE International Conference on Robotics and Automation*, Stockholm, Sweden, 2016, pp. 2082-2087.

[6]. J. Koh, E. Yang, G. Jung, Jumping on water: Surface tension-dominated jumping of water striders and robotic insects, *Science*, Vol. 349, 2015, pp. 517-521.

[7]. Y. Bai, J. B. Snyder, M. A. Peshkin, M. A. MacIver, Finding and identifying simple objects underwater with active electrosense, *International Journal of Robotics Research*, Vol. 34, 2015, pp. 1255-1277.

[8]. J. Kalaska, From intention to action: motor cortex and the control of reaching movements, *Advances in Experimental Medicine and Biology*, Vol. 629, 2009, pp. 139-178.

[9]. B. Bruno, F. Mastrogiovanni, A. Sgorbissa, Functional requirements and design issues for a socially assistive robot for elderly people with mild cognitive impairments, in *Proceedings of the 22nd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'13)*, 2013, pp. 768-773.

[10]. A. Campeau-Lecours, U. Côté-Allard, D. Vu, F. Routhier, B. Gosselin, C. Gosselin, Intuitive adaptive orientation control for enhanced human-robot interaction, *IEEE Transactions on Robotics*, Vol. 35, Issue 2, 2019, pp. 509-520.

[11]. C. F. Cadieu, et al., Deep neural networks rival the representation of primate IT cortex for core visual object recognition, *PLoS Computational Biology*, Vol. 10, 2014, e1003963.

[12]. M. Churchland, J. Cunningham, M. Kaufman, Neural population dynamics during reaching, *Nature*, Vol. 487, 2012, pp. 51-56.

[13]. E. Cross, R. Hortensius, A. Wykowska, From social brains to social robots: applying neurocognitive insights to human-robot interaction, *Philosophical Transaction of Royal Society B*, Vol. 374, 2019, 20180024.

[14]. J. Diedrichsen, R. Shadmehr, R. B. Ivry, The coordination of movement: optimal feedback control and beyond, *Trends in Cognitive Sciences*, Vol. 14, 2010, pp. 31-39.

[15]. P. Davidson, D. Wolpert, Motor learning and prediction in a variable environment, *Current Opinion in Neurobiology*, Vol. 13, Issue 2, 2003, pp. 232-237.

[16]. M. Erden, A. Billard, Robotic assistance by impedance compensation for hand movements while manual welding, *IEEE Transactions on Cybernetics*, Vol. 46, Issue 11, 2016, pp. 2459-2472.

[17]. F. Flacco, T. Kröger, A. Luca, O. Khatib, A depth space approach to human-robot collision avoidance, in *Proceedings of the IEEE International Conference on Robotic and Automation (ICRA'12)*, 2012, pp. 338-345.

[18]. W. Fang, F. Chao, L. Yang, et al., A recurrent emotional CMAC neural network controller for vision-based mobile robots, *Neurocomputing*, Vol. 334, 2019, pp. 227-238.

[19]. R. Gonzalez, M. Fiacchini, K. Iagnemma, Slippage prediction for off-road mobile robots via machine learning regression and proprioceptive sensing, *Robotics and Autonomous Systems*, Vol. 105, 2018, pp. 85-93.

[20]. D. Hassabis, D. Kumaran, C. Summerfield, M. Botvinick, Neuroscience-inspired artificial intelligence, *Neuron*, Vol. 95, 2017, pp. 245-258.

[21]. C. Huijnen, M. Lexis, L. Witte, Matching robot KASPAR to autism spectrum disorder (ASD) therapy and educational goals, *International Journal of Social Robotics*, Vol. 8, Issue 4, 2016, pp. 445-455.

[22]. Y. Y. Jia, J. Zhao, M. Alfatlawi, N. Xi, Brain-like control and adaptation for intelligent robots, in Brain inspired intelligent robotics: The intersection of robotics and neuroscience, *Science/AAAS*, 2016, pp. 20-24.

[23]. L. Koumakis, C. Chatzaki, E. Kazantzaki, E. Maniadi, M. Tsiknakis, Dementia care frameworks and assistive technologies for their implementation: A review*, IEEE*

Reviews in Biomedical Engineering, Vol. 12, 2019, pp. 4-18.

[24]. N. Kriegeskorte, P. K. Douglas, Cognitive computational neuroscience, *Nature Neuroscience*, Vol. 21, 2018, pp. 1148-1160.

[25]. N. Ludolph, M. Giese, W. Llg, Interacting learning processes during skill acquisition: learning to control with gradually changing system dynamics, *Scientific Reports*, Vol. 7, 2017, 325.

[26]. B. Lake, T. Ullman, J. Tenenbaum, S. Gershman, Building machines that learn and think like people, *Behavioral and Brain Sciences*, Vol. 40, 2017, e253.

[27]. L. Brenner, Marriage of mind and machine, *Nature*, Vol. 571, 2019, pp. S15-S17.

[28]. S. Lemaignan, M. Warnier, E. Sisbot, A. Clodic, R. Alami, Artificial cognition for social human-robot interaction: An implementation, *Artificial Intelligence*, Vol. 247, 2017, pp. 45-69.

[29]. E. Magrini, F. Flacco, A. Luca, Control of generalized contact motion and force in physical human-robot interaction, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'15)*, 2015, pp. 2298-2304.

[30]. A. H. Marblestone, G. Wayne, K. P. Kording, Towards an integration of deep learning and neuroscience, *Frontiers in Computational Neuroscience*, Vol. 10, 2016, 94.

[31]. A. Meguenani, V. Padois, P. Bidaud, Control of robots sharing their workspace with humans: An energetic approach to safety, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'15)*, 2015, pp. 4678-4684.

[32]. V. Mnih, et al., Human-level control through deep reinforcement learning, *Nature*, Vol. 518, 2015, pp. 529-533.

[33]. F. Niroui, K. Zhang, Z. Kashina, G. Nejat, Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments, *IEEE Robotics and Automation Letters*, Vol. 4, 2019, pp. 610-617.

[34]. S. Meyer, C. Fricke, Robotic companions in stroke therapy: A user study on the efficacy of assistive robotics among 30 patients in neurological rehabilitation, in *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'17)*, 2017, pp. 135-142.

[35]. E. Robert, The Art and Science of Java, *Pearson*, 2008.

[36]. P. R. Roelfsema, A. van Ooyen, Attention-gated reinforcement learning of internal representations for classification, *Neural Computation*, Vol. 17, 2005, pp. 2176-2214.

[37]. S. Nandhakumar, V. Muthukumaran, K. Prakash, V. Shunmughanaathan, Position control of industrial robotic manipulator using variable structure control system with single term Haar wavelet series method, *Journal of Vibration and Control*, Vol. 21, Issue 12, 2015, pp. 2465-2483.

[38]. S. Rabbitt, A. Kazdin, B. Scassellati, Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use, *Clinical Psychology Review*, Vol. 35, 2015, 35.

[39]. D. L. K. Yamins, et al., Performance-optimized hierarchical models predict neural responses in higher visual cortex, *Proceedings of the National Academy of Sciences*, Vol. 111, 2014, pp. 8619-8624.

[40]. I. Rodriguez, A. Astigarraga, E. Lazkano, J. M. Martínez-Otzeta, I. Mendialdua, Robots on stage: A cognitive framework for socially interacting robots, *Biologically Inspired Cognitive Architectures*, Vol. 25, 2018, pp. 17-25.

[41]. U. Rokni, H. Sompolinsky, How the brain generates movement, *Neural Computation*, Vol. 24, 2011, pp. 289-331.

[42]. C. J. Spoerer, P. McClure, N. Kriegeskorte, Recurrent convolutional neural networks: A better model of biological object recognition, *Frontiers in Psychology*, Vol. 8, 2017.

[43]. J. Slotine, W. Li, On the adaptive control of robot manipulators, *International Journal of Robotics Research*, Vol. 6, 1987, pp. 49-59.

[44]. J. Schmidtler, K. Bengler, F. Dimeas, A. Campeau-Lecours, A questionnaire for the evaluation of physical assistive devices (QUEAD): Testing usability and acceptance in physical human-robot interaction, in *Proceedings of the IEEE International Conference on Systems Man and Cybernetics (SMC'17)*, 2017, pp. 876-881.

[45]. J. A. Reggia, G. E. Katz, G. P. Davis, Humanoid cognitive robots that learn by imitating: Implications for consciousness studies, *Frontiers in Robotics and AI*, Vol. 5, Issue 1, 2018.

[46]. S. Schneider, M. Goerlich, F. Kummer, A framework for designing socially assistive robot interactions, *Cognitive Systems Research*, Vol. 43, 2017, pp. 301-312.

[47]. A. Zanchettin, N. Ceriani, P. Rocco, H. Ding, B. Matthias, Safety in human-robot collaborative manufacturing environments: Metrics, control, *IEEE Transaction on Automation Science and Engineering*, Vol. 13, Issue 2, 2016, pp. 882-893.

[48]. D. Zhang, B. Wei, On the development of learning control for robotic manipulators, *Robotics*, Vol. 6, 2017, 23.

[49]. M. Sun, H. Li, Y. Li, A unified design approach to repetitive learning control for systems subject to fractional uncertainties, *ASME Journal of Dynamic System, Measurement and Control*, Vol. 140, Issue 6, 2018, 061003.

[50]. P. Shapshak, Artificial Intelligence and brain, *Bioinformation*, Vol. 14, Issue 1, 2018, pp. 38-41.

[51]. J. Wainer, B. Robins, F. Amirabdollahian, K. Dautenhahn, Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism, *IEEE Transactions on Autonomous Mental Development*, Vol. 6, Issue 3, 2014, pp. 183-199.

[52]. T. Yamawaki, H. Ishikawa, M. Yashima, Iterative learning of variable impedance control for human-robot cooperation, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16)*, 2016, pp. 839-844.

[53]. D. L. K. Yamins, et al., Performance-optimized hierarchical models predict neural responses in higher visual cortex, *Proceedings of National Academy of Sciences*, Vol. 111, Issue 23, 2014, pp. 8619-8624.

[54]. M. Yu, C. Li, Robust adaptive iterative learning control for discrete-time nonlinear systems with time-iteration-varying parameters, *IEEE Transactions on Systems Man and Cybernetics: Systems*, Vol. 47, Issue 7, 2017, pp. 1737-1745.

(009)

Learning from Demonstration for Collaborative Robots

A. Rizzotti-Kaddouri [1], M. Kunze [2], L. Jeanneret [1], L. Depierraz [2] and N. Ouerhani [1]
[1] Haute Ecole Arc Ingénierie, School of Engineering, Neuchâtel, Switzerland
[2] Haute Ecole d'Ingénierie et de Gestion du Canton de Vaud, Yverdon-les-Bains, Switzerland
aicha.rizzotti@he-arc.ch

Summary: This article presents our work from research project which objective is to allow a robot to perform pick & place and assembly tasks by intuitively teaching and programming the robot trajectories from human demonstrations. Based on motion acquisition systems, we aim at developing a system capable of acquiring and analyzing the manipulation actions performed by an operator to extract their primitives and compound characteristics. A Leap Motion sensor and a 3D camera are used to acquire gestures and movements at different scales and objects position. Classification algorithms and deep learning models are used in order to recognize gestures. An expert system is allowing the translation of recognized gestures into robot trajectories. The challenge in our case is the automation of tasks through artificial intelligence. ABB's YuMi Robot is used to validate our solution.

Keywords: Observational learning, Collaborative robots, Robot learning, Cognitive systems, Cyber-physical systems, Imitation learning, Gesture and movement capture.

1. Introduction

Robot learning by demonstration (LfD) is a promising approach that allows the transfer of robotic solution prototypes from the academic world to the business world [1].Technological advances in recent years have allowed robots to get out of their cages and work side by side with human operators. With these collaborative robots, a new era of industrial robotics is opening. The potential market for collaborative robots is expected to grow from $ 400 million in 2017 to $ 7.5 billion in 2027. Brought to work with and alongside humans, these robots will have to provide advanced skills. Classic robot programming (complex and time consuming) is reaching its limits in the face of this new challenge. Effort is, therefore, being invested to reduce programming time and reduce the cost of deploying a robotic system by up to 60 % according to the study [2].

The goal of our project LD4Robots (Learning from Demonstration for collaborative Robots) is to design and prototype a learning system that will be used to automatically program manipulation tasks of an industrial collaborative robot by reproducing the actions demonstrated by an operator (see Fig. 1).

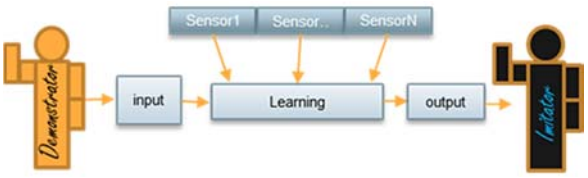


Fig. 1. LDF approach.

In this project, we propose to apply this method first to a "Pick & Place" [5] use case and in a second step to an assembly one. Currently, we are working on the first use case, which will be the basis for developing the second one. In our project, we favored an approach for LFD based on a vision system consisting of a Leap Motion [3] and a RealSence camera unlike kinesthetic guidance [4] which refers to manual movements of the robot by haptic interaction through haptic sensors. Note that the second approach simplifies the task of correspondence learning but on the other hand requires a robot allowing haptic interaction.

It should be emphasized that our workflow phases are:

- Data capture and fusion;
- Learning phase;
- Reproduction by the robot of the movement or the learned action.

2. Architecture and Set-up

The overall architecture of our solution is depicted in Fig. 2. The solution requires the integration of different components running on different Operating Systems (OS). In order to solve OS compatibility issues, we have created a specific architecture: an application running on Windows OS captures the movements of an operator using different sensors (LeapMotion, RealSense), assembles the data and saves the movement of the operator in a file located on a server. Once the demonstration is done by the operator, the data collected by the sensors, the trajectories and gestures classified saved in a file, the goal is then to program the YuMi robot and reproduce the task. To do it the trajectories are computed and programmed by MoveIt! running on ROS (Robot Operating System) a motion planning framework and then communicated to the robot using the EGM (Externally Guided Motion) library provided by ABB. This second part is running on a Linux platform.

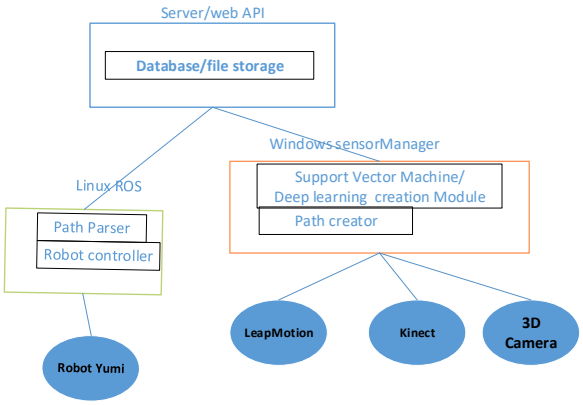


Fig. 2. Architecture.

The set-up on which the taught demonstration is reproduced is composed of an ABB YuMi robot (dual arm robot; each arm having 7 DoF), a Leap Motion sensor and an Intel RealSense camera (see Fig. 3).

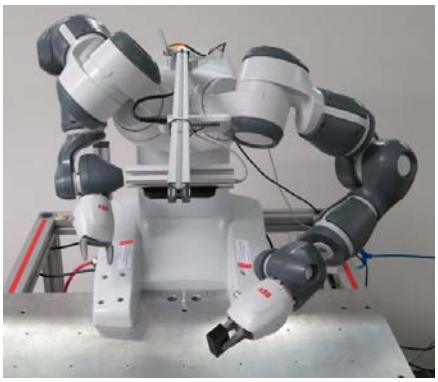


Fig. 3. ABB YuMi robot, Leap Motion and RealSense sensors.

3. Methodology

Our approach is composed of two different parts: teaching and reproduction. For the "pick & place" scenario, three essential types of gestures are initially considered:

- Movement: this is an action which recognition and precise description will guide the trajectory of the robotic arm;
- Pick: action of picking an object in a "Pick & Place" scenario, one of the most frequent scenarios in collaborative robotics. Determining when and how to grab an object is an essential capability;
- Place: action of placing an object. This action makes sense and usually follows the picking of an object.

For the teaching part we use the 3D camera to track the position and orientation of the object before picking it and after placing it. On the other, hand the leap motion is used to track the fine movements and track different fine gestures such as the pinch gesture. This latter gives also the start of the recording of the trajectory of the hand and the detection of the final new position of the object.

Recognizing these three actions requires the automatic interpretation of the data provided by the different sensors. We use SVM (Support Vector Machine) classifiers to differentiate the gestures and we are currently working on another version using deep learning to compare the performance of the two ways of doing.

Regarding the reproduction of the trajectory, it is therefore a matter of matching the trajectory described by human movements with the typical trajectories of a robotic arm. For the moment, we extract from the trajectories given by the Leap Motion 4 poses (position and orientation) of the hand of the demonstrator that are the starting point, the "picking" point, the "placing" point and the end point. These 4 points enable us to fully describe and programme the trajectory. In a second step we will also include obstacle avoidance by closely tracking the movement of the hand of the operator between the picking point and the placing point during the movement.

4. Conclusions

So far, thanks to optical sensors (LeapMotion and RealSense), we are able to capture, classify, recognize and reproduce all movements and gestures involved into a simple "Pick & place" use case. The RealSense camera enables us to improve the accuracy obtained with the Leap Motion sensor when it is needed (to get the exact picking and placing positions). The next steps will be to include obstacle avoidance and also to perform more complex tasks like assembly tasks.

Acknowledgements

This project is funded by the University of Applied Sciences and Arts of Western Switzerland.

References

[1]. B. D. Argall, S. Chernova, M. Veloso, B. Browning, A survey of robot learning from demonstration, *Robotics and Autonomous Systems*, Vol. 57, Issue 5, 2009, pp. 469-483.

[2]. J. A. Marvel, et al., Tools for robotics in SME work cells: challenges and approaches for calibration and registration, Interagency/Internal Report (NISTIR)-8093, *NIST*, 2015.

[3]. L. Shao, Hand movement and gesture recognition using Leap Motion Controller, Virtual Reality, Course Report, *Stanford*, 2016.

[4]. Z. Zhu, H. Hu, D. Gu, Robot performing peg-in-hole operations by learning from human demonstration, in *Proceedings of the 10th Computer Science and Electronic Engineering (CEEC'18)*, 2018, pp. 30-35.

[5]. A. Sena, M. Howard, Quantifying teaching behavior in robot learning from demonstration, *The International Journal of Robotics Research*, Vol. 39, Issue 1, 2020, pp. 54-72.

(010)

Model-driven Engineering of Gateways for Industrial Automation

P. Denzler [1], D. Ramsauer [1] and W. Kastner [1]
[1] TU Wien, Automation Systems Group, Treitlstr. 1-3/4., 1040 Vienna, Austria
Tel.: +43 1 58801-18311, fax: +43 1 58801-18391
E-mail: patrick.denzler@tuwien.ac.at

Summary: Middleware gateways are a feasible solution for improving interoperability between legacy connectivity technologies in industrial automation and enterprise information technology. However, developing, configuring, and deploying gateways on different devices for connecting middlewares are not well examined. A proposed gateway generation process that builds upon a generic gateway model in the Architecture Analysis & Design Language (AADL) and allows platform-specific code generation addresses those challenges. Additional, configuration files define all necessary bindings. The paper concludes by outlining subsequent steps by pointing out research challenges in the presented process.

Keywords: Middleware, Gateways, Industrial automation, Industrial Internet of Things (IIoT), Architecture Analysis & Design Language (AADL).

1. Introduction

Accessing information in industrial automation has always been a critical element for a system's optimal functioning, but often hindered by the used architecture [1]. From a historical perspective, typical automation systems follow a hierarchical architecture often visualised as an automation pyramid (Fig. 1) [2]. The first two levels represent sensors and actuators, followed by controllers (e.g., programmable logic controllers (PLCs)) which are responsible for the interaction with the underlying technical process and the processing of automation functions (e.g., open/closed-loop control), respectively. With the third SCADA systems level, the lower levels represent the operational technology (OT) part of automation [3].

Compared to the two top levels, summarised as information technology (IT), OT has more stringent requirements (e.g., on real-time and safety) achieved by automation and industrial communication technology. On the other hand, IT uses off-the-shelf components and communication systems ready for use with the Internet Protocol (IP) suite part of any enterprise / IT environment. Typical applications in the upper levels are Manufacturing Execution Systems (MES) and Enterprise-Resource-Planning (ERP) tool suites that process and abstract the data from the lower levels. Accessing OT information is complicated and is often referred to as the OT/IT gap [2].

Given more recent initiatives, e.g., the Industrial Internet of Things (IIoT) or Industry 4.0, access to information is becoming even more critical. IIoT aims to close the OT/IT gap by introducing the same IP technology at all levels, thus creating a more flat and transparent architecture. Such a system design would make it possible to address a single sensor on the shop floor directly from a top-level application (e.g., located in the cloud).

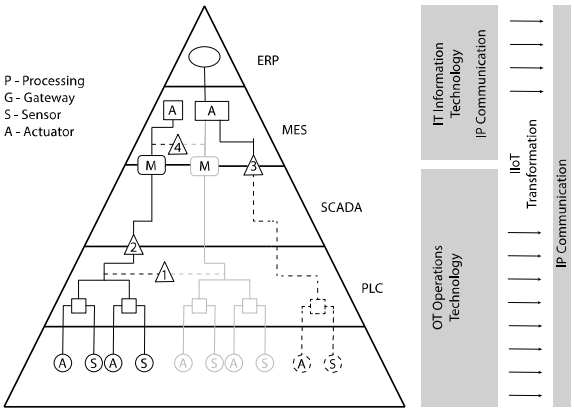


Fig. 1. OT / IT gap and gateway positions in the Automation Pyramid.

New technological developments such as fog and cloud computing or time-sensitive networking (TSN) further promise to close the OT/IT gap [4].

An often forgotten problem in the long history of industrial communication is that the domain looks back on several decades of industrial communication technology. Due to long life cycles, it is not uncommon to find relatively old, still correctly working systems [5]. Therefore, a sudden change towards an all-new IP architecture is not very realistic. There is, therefore, a need for interim solutions, for example, gateways to connect older OT protocols with newer technology and Internet protocols of the IT levels [6].

In this context, this article describes the possible positions of gateways within the automation pyramid. It presents arguments why middleware gateways are a feasible solution to close the OT/IT gap. It also describes the related challenges, followed by a process that addresses them. The results form the basis for further research about middleware gateways for industrial automation.

Section 2 presents possible positions of gateways in the automation pyramid and the historical background. Section 3 introduces the middleware gateway concept and lays the ground for the related challenges in Section 4. Within Section 5, the gateway generation process is introduced and discussed in Section 6. Section 7 concludes the paper and provides an outlook on future work.

2. Gateways in the Automation Pyramid

Gateways have always been part of industrial automation. They found their application at different automation pyramid levels by connecting protocols horizontally or vertically with IT applications [1].

In the early days of automation when fieldbus systems replaced parallel and discrete cabling in the lower levels of the automation pyramid, early suggestions appeared to improve interoperability with gateways [7]. The focus of these dedicated automation networks (e.g. Controller Area Network (CAN), PROFIBUS or INTERBUS) however was to overcome the restrictions of parallel cabling and not on interoperability [7]. Therefore, low-level horizontal gateways, as depicted in Fig. 1, remained an exception.

With the advent of the Internet, the relevant technology also made its way into the automation domains. Due to the uniform Ethernet standard, gateways seem to be a relic of the past. However, Ethernet lacked actual real-time capabilities, which led to the development of several new dedicated protocols and reintroduced the need for gateways.

Protocols such as PROFINET or EtherCAT either use parts of the OSI layers or adapt them to achieve low latency communication. In contrast to Internet technology, such protocols are OT specific and not intended for communication with the IT domain [8]. The specific real-time requirements make creating horizontal gateways a problematic endeavour. Research has shown that directly connecting such protocols can create serious security risks [1]. Nowadays, the new initiative creating time-sensitive networks (TSN) opens up new possibilities [4].

Vertical gateways are another solution to overcome OT and IT's separation as they transfer automation data to the IP-based networks. The gateways abstract a certain amount of data and functions from OT to the enterprise environment [9]. However, creating vertical gateways requires extensive knowledge of the protocol and the connected application (MES, ERP system). The uniqueness of each gateway makes it challenging to maintain and reuse or adjust it for other applications. One possible solution is to connect vertical gateways with middleware that is better suited to the IT environment.

With middleware, IT applications can access the OT world's data points with the necessary abstraction and additional information. There are different ways to access legacy protocols. Either directly with specific APIs and drivers or as mentioned via vertical gateways. Since middlewares offer standardised access to all IT applications, their use would reduce the number of required gateways [6]. However, the number of middleware and protocol combinations is still considerably large and would hinder interoperability.

At this point, middleware gateways could improve the situation by building bridges between middlewares. The complexity of such gateways is high because middlewares sometimes use different communication patterns, e.g., OPC Unified Architecture (OPC UA) (Server-Client), MQTT (Broker) or Data Distribution Service (DDS) (publish-subscribe) and means of transport (TCP/IP, UDP, HTTP or CoAP) [9]. Nonetheless, the IT compatibility allows the flexible deployment, e.g., on fog nodes or in the cloud and are therefore a suitable solution to play an essential role in closing the OT/IT gap.

3. Middleware Gateway Concept

Middleware gateways could reduce the total amount of required gateways in IIoT. The goal is to connect the existing middlewares, which are either already attached to legacy protocols or provide drivers or APIs. Such a proposal is not new. The Industrial Internet Consortium (IIC) published a related study to identify the leading middleware technologies for their suitability for IIoT applications [10]. After an in-depth evaluation OPC UA, DDS, oneM2M and Web Services turned out to be best candidates for all types of industrial sectors. Connecting those "main" middlewares with gateways would significantly improve interoperability. Secondary gateways would connect other middlewares or legacy protocols to one of the main middlewares (cf. Fig. 2).

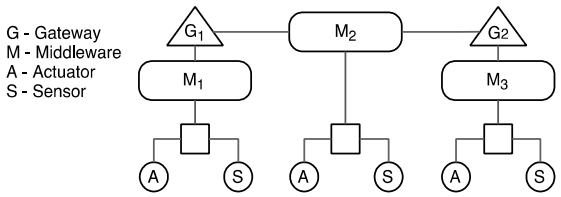


Fig. 2. Main middleware gateways concept.

As part of the IIC, the OMG Group published an OPC UA / DDS gateway specification [11] to follow up on the concept. Since OPC UA is a server-client-based middleware and DDS follows the publish-subscribe paradigm, the specification addresses these differences by utilising configurable bi-directional bridges. Since its publication in 2019, the gateway specification received some attention in the industry and led to a proprietary implementation. In academic research, the focus remains mainly on the gateway's information exchange and functionality aspects by implementing reduced prototypes as in [12]. However, none of the studies focused on the challenges associated with implementing, configuring, and deploying such gateways.

4. Challenges Related to Middleware Gateways

A recent use-case study about fog computing in automation systems to consolidate SCADA functionality, and using an OPC UA / DDS gateway prototype revealed several challenges [13]. In this use case, DDS acted as a bridge between two OPC UA networks and provided remote maintenance functionality. The trials revealed mostly problems related to the implementation of the gateway prototype itself.

One other significant difficulty is the configuration of the gateway. Configuration requires in-depth knowledge about both networks; i.e., the node addresses, the data to be transferred or access rights. These hurdles directly impact the gateway's deployment and scalability because some information must be in place before the compilation. If such dependencies exist, the dynamic deployment of a gateway on different resources, for example on fog nodes, would require an automated compile process or a change in the gateway architecture. Other studies report similar problems with gateways, despite different protocols and architectures [14]. For example, Dionisio et al. [15] present an OPC UA / MQTT gateway based on Node-RED for "loom machines" where they use a static gateway configuration. Some authors [16] try to counter the configuration issue with a software-defined perspective or with a particular additional middleware [17].

Other problems encountered in the use case are related to security and latency issues. An additional firewall that protects the gateway from malicious access solved the problem. However, this solution required additional configuration effort. It would have been more user-friendly if the gateway had supported the security functions used by DDS and OPC UA. While latency and overhead were not a critical problem in the use case, some considerations had to be made. Other authors report similar problems and use HTTP to bridge protocols to reduce overhead and latency [18].

In summary, to enable middleware gateways to bridge the OT/IT gap and connect legacy protocols, those challenges need to be addressed first.

5. Gateway Generation Process

Gateway configuration, deployment and scalability are the most relevant challenges that a potential solution needs to address. The suggested process foresees using a modelling language to model the different aspects of a gateway and code generation to achieve higher flexibility. Potential candidates are Unified Modeling Language (UML), Systems Modeling Language (SysML) and the Architecture Analysis & Design Language (AADL). All languages support the modelling of complex applications and allow code generation. However, only SysML and AADL are capable of specifying hardware and communication. AADL originating out of avionics provides analysis possibilities for embedded, real-time systems, supporting its use in industrial automation [20].

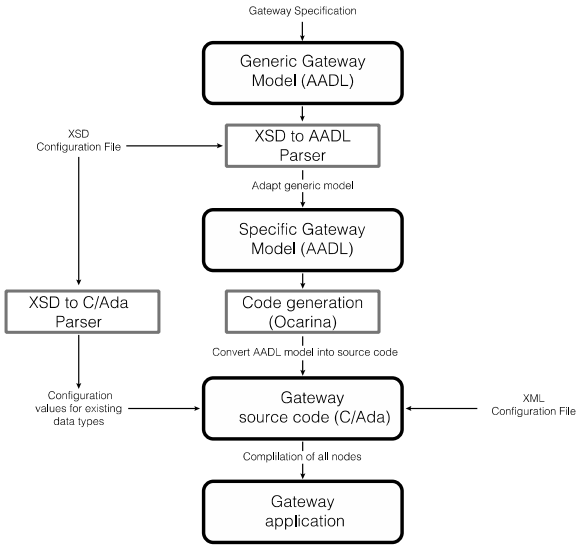


Fig. 3. Gateway generation process.

The suggested process assumes the usage of AADL and related tools. However, it is also applicable to other modelling languages. The first step is the creation of a generic gateway model based on a specification. This unique specification depends on the middleware combination and defines necessary bindings and features. The created general model is the foundation for the further code generation and acts as an empty shell for specific gateway instances. For the actual code generation, each gateway instance requires further "hardcoded" information, e.g., data type mappings or node addresses. The proposed process foresees two stages for adding the required information. In the first stage, a configuration file provides the parser everything to create a specific gateway model. The following code generation then produces code. In AADL, the Ocarina toolchain creates either C or ADA code out of the model [20]. It might be worth to mention that the code generation is dependent on the chosen hardware environment, e.g., an embedded system. This feature is essential when considering that gateways might run on resource-restricted edge devices or fog nodes. In the last step before compiling the code, an XML configuration file provides the final mappings that end the gateway creation process.

6. Discussion

Although the proposed process is relatively simple, it addresses the most relevant challenges related to middleware development and use of gateways. Adopting a general gateway model allows the

generation of specific and compiled gateway instances and reduces the configuration effort when adjustments are required. A weak point of the process is that the model entirely depends on a given gateway specification unique for each middleware combination; however, following the main gateway idea with only a few combinations reduces the effort.

As other studies indicated, the configuration of the gateway is a core challenge. Despite the proposed approach to integrating the configuration into the building process, the required configuration files are complex. This fact limits the flexibility as the configuration is static. A possible solution would be using an external tool that creates the files. In the best case, the tool obtains the needed (network) information from both middlewares automatically. OPC UA, for example, applies information models to structure all involved nodes. Another option is to change the architecture of the gateway, that would allow dynamical changes during runtime. Such functionality, however, might increase the complexity and is not supported in all programming languages.

If the configuration is static, the process will allow an automated generation of gateways. For instance, applied in fog computing, a deployment tool creates new instances when one gateway cannot handle all traffic. Moreover, though the process aims for middleware gateways, it is applicable for direct vertical gateways to connect legacy protocols with IT applications.

6. Conclusion

The paper presents a process to address challenges related to the development and deployment of middleware gateways which are an efficient way of connecting legacy OT protocols to the IT environment. A previously conducted use case study and related work indicate that the most relevant challenges are configuration, deployment, scalability and security. The model-based approach includes a two-step configuration, code generation for different platforms, and automation potential. Further studies will include the actual modelling of an OPC UA / DDS gateway with the proposed AADL language.

Acknowledgements

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 764785, FOR A – Fog Computing for Robotics and Industrial Automation.

References

[1]. T. Sauter, M. Lobashov, How to access factory floor information using internet technologies and gateways, *IEEE Transactions on Industrial Informatics*, Vol. 7, Issue 4, Nov. 2011, pp. 699-712.

[2]. D. R. Harp, G. Bengt, IT/OT Convergence Bridging the Divide, *NexDefense*, 2015.

[3]. M. Wollschlaeger, T. Sauter, J. Jasperneite, The future of industrial communication: automation networks in the era of the Internet of Things and Industry 4.0, *IEEE Industrial Electronics Magazine*, Vol. 11, Issue 1, March 2017, pp. 17-27.

[4]. M. Aazam, S. Zeadally, K. A. Harras, Deploying fog computing in industrial Internet of Things and Industry 4.0, *IEEE Transactions on Industrial Informatics*, Vol. 14, Issue 10, Oct. 2018, pp. 4674-4682.

[5]. K. F. Tsang, M. Gidlund, J. Åkerberg, Guest editorial industrial wireless networks: Applications, challenges, and future directions, *IEEE Transactions on Industrial Informatics*, Vol. 12, Issue 2, April 2016, pp. 755-757.

[7]. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of Things: A survey on enabling technologies, protocols, and applications, *IEEE Communications Surveys & Tutorials*, Vol. 17, Issue 4, 2015, pp. 2347-2376.

[8]. G. Wood, Survey of LANs and standards*, Computer Standards & Interfaces,* Vol. 6, Issue 1, 1987, pp. 27-36.

[9]. P. Danielis, *et al*., Survey on real-time communication via Ethernet in industrial automation environments, in *Proceedings of the IEEE International Conference on Emerging Technology and Factory Automation (ETFA'14)*, Barcelona, 2014, pp. 1-8.

[10]. T. Sauter, The continuing evolution of integration in manufacturing automation, *IEEE Industrial Electronics Magazine*, Vol. 1, Issue 1, 2007, pp. 10-19.

[11]. Industrial Internet Reference Architecture (IIRA), *Industrial Internet Consortium*, 2015.

[12]. OPC UA/DDS Gateway Version 1.0. (2020, October 15), Object Management Group, https:// www.omg.org/spec/ DDS-OPCUA/1.0

[13]. R. Endeley, T. Fleming, N. Jin, G. Fehringer, S. Cammish, A smart gateway enabling OPC UA and DDS interoperability, in *Proceedings of the IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI'19)*, 2019, pp. 88-93.

[14]. P. Denzler, J. Ruh, M. Kadar, C. Avasalcai, W. Kastner, Towards consolidating industrial use cases on a common fog computing platform, in *Proceedings of the 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'20)*, Vienna, Austria, 2020, pp. 172-179.

[15]. R. R. Peña, R. D. Fernández, M. Lorenc, A. Cadiboni, Gateway OPC UA/Modbus applied to an energy recovery system identification, in *Proceedings of the XVIII Workshop on Information Processing and Control (RPIC'19)*, Bahía Blanca, Argentina, 2019, pp. 235-240.

[16]. R. Dionisio, S. Malhao, P. Torres, Development of a smart gateway for a label loom machine using industrial IoT technologies, *International Journal of Online and Biomedical Engineering (IJOE)*, Vol. 16, Issue 4, 2020, pp. 6-14.

[17]. Z. Jiang, Y. Chang, X. Liu, Design of software-defined gateway for industrial interconnection, *Journal of Industrial Information Integration*, Vol. 18, 2020, 100130.

[18]. A. Ioana, A. Korodi, VSOMEIP – OPC UA gateway solution for the automotive industry, in *Proceedings of the IEEE International Conference on Engineering,*

Technology and Innovation (ICE/ITMC'19), Valbonne Sophia-Antipolis, France, 2019, pp. 1-6.

[19]. M. A. A. da Cruz, J. J. P. C. Rodrigues, P. Lorenz, P. Solic, J. Al-Muhtadi, V. H. C. Albuquerque, A proposal for bridging application layer protocols to HTTP on IoT solutions, *Future Generation Computer Systems*, Vol. 97, 2019, pp. 145-152.

[20]. P. H. Feiler, D. P. Gluch, J. J. Hudak, The Architecture Analysis & Design Language (AADL): An Introduction, Technical Report, Software Engineering Inst., *Carnegie-Mellon Univ*., Pittsburgh, PA, 2006.

(011)

Condition Monitoring of Drive Trains by Data Fusion of Acoustic Emission and Vibration Sensors

Oliver Mey [1], André Schneider [1], Olaf Enge-Rosenblatt [1], Dirk Mayer [1], Christian Schmidt [2], Samuel Klein [2] and Hans-Georg Herrmann [2]
[1] Fraunhofer Institute of Integrated Circuits IIS, Division Engineering of Adaptive Systems, Dresden, Germany
[2] Fraunhofer IZFP Institute for Nondestructive Testing, Saarbrücken, Germany
Tel.: +49 351 4640795
E-mail: oliver.mey@eas.iis.fraunhofer.de

Summary: Early damage detection and classification by condition monitoring systems is crucial to enable predictive maintenance of manufacturing systems and industrial facilities. The data analysis can be improved by applying machine learning algorithms and fusion of data from heterogeneous sensors. This paper presents an approach for a step-wise integration of classifications gained from vibration and acoustic emission sensors, in order to combine the information from signals acquired in the low and high frequency range. A test rig comprising a drive train and bearings with small artificial damages is used for acquisition of experimental data. The results indicate that an improvement of damage classification can be obtained using the proposed algorithm of combining classifiers for vibrations and acoustic emission.

Keywords: Condition monitoring, Vibration, Acoustic emission, Drive train, Data fusion, Machine learning.

1. Introduction

Increasing demand for efficient, reliable and available industrial production systems has led to the development of systems for continuous condition monitoring (CM). By the analysis of sensor signals, failures and wear can be detected in an early stage, enabling cost-effective predictive maintenance and prohibiting severe damages [1-3].

2. Applications and Methods for Condition Monitoring

A widely implemented application is condition monitoring of bearings. These are critical parts of many industrial drives. Since bearings exist in a wide range of geometries and are driven at different speeds and loads, CM systems have to be adapted individually for a sufficient damage detection capability [4].

A common monitoring approach includes vibration measurements, usually conducted with accelerometers. It has been successfully applied for the detection of mechanical faults like imbalances [2]. Still, the performance of vibration-based CM systems suffers when dynamic acceleration levels become low due to slow rotation speeds, when disturbance levels are high due to coupling with noisy gearboxes or other drivetrain components, and in case of varying speeds.

CM based on ultrasonic structure-borne noise or acoustic emission has been proven to be an alternative in some cases [5]. The method is based on the detection of impulsive events (bursts) sent out from the rolling contact of damaged bearing elements. While being very sensitive, disadvantages of the method include the limited range of the impulsive waves in the mechanical structure and being insensitive to global mechanical faults like misalignment of drive shafts, imbalances or loose parts.

For both methods, vibration approach and ultrasonic approach, a number of signal analysis algorithms have been developed, either in time or frequency domain. More recent approaches use the corresponding sensor data as input to powerful machine learning algorithms.

3. Data Fusion Approach

In this paper, a combined system is presented, which utilizes low-frequency vibration signals from accelerometers and high-frequency acoustic emission (AE) signals from ultrasonic transducers. By feeding the signals into a multi-stage classification algorithm (see Fig. 1), the capabilities of the detection should be enhanced to a broad range of faults and damages. The classifier architecture thereby consists of 3 parts, which all are trained to detect the defect state of the system:

- A classifier which receives data from the acoustic emission sensor (AE classifier);
- A classifier which receives data from the acceleration sensor (vibration classifier);
- A classifier which receives activations of the AE and vibration classifiers at intermediate layers (combined classifier).

The intension *for this approach* was to increase the sensitivity and to improve the resilience with respect to sensor failures. The proposed classifier architecture provides classifications based only on single sensor streams on the one hand (using the data from one sensor), and classifications based on fused sensor information on the other hand (using all available sensors). Thereby, the sensitivity of the respective

sensor data for the detection of different defect states can be evaluated and the performance gain due to the conducted sensor fusion can be assessed. Additionally, faults can even be detected if a single sensor stream is missing, albeit not with the maximum prediction accuracy.

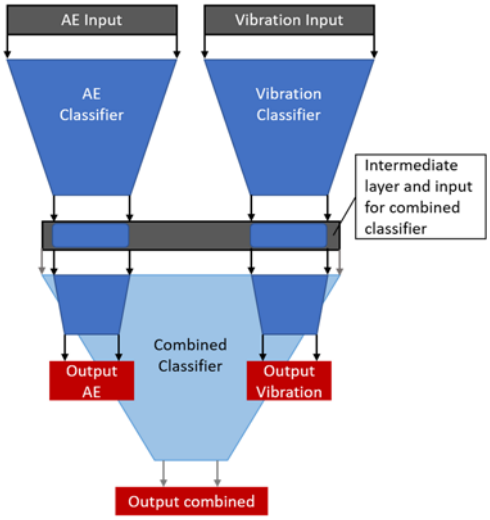


Fig. 1. Sketch of the proposed classifier architecture.

4. Laboratory Set Up

The proposed approach is demonstrated with a drive train set up, at which multiple bearing failures can be simulated (see Fig. 2).

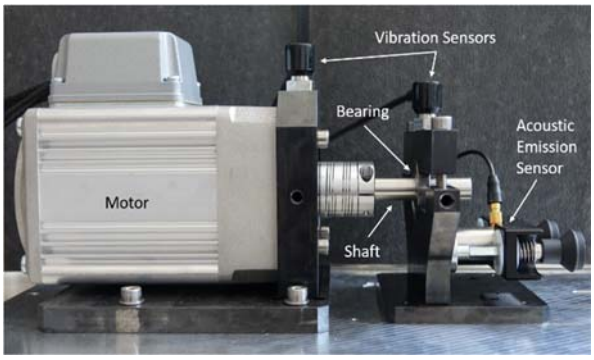


Fig. 2. Drive train set up with vibration and acoustic emission sensors.

The basic concept documented in [2] is extended by an instrumentation with AE sensors. The set up used here comprises a motor (WEG V3.5111) connected to a bearing shaft by a flexible coupling. The bearing is a standard roller bearing type E12.

For vibration measurements, the setup is instrumented with ICP accelerometers (PCB 607A11, 100 mV/g) at the bearings of the motor and the shaft. AE signals are acquired by a piezo transducer (Vallen VS30V, 20 kHz – 80 kHz) connected to an AEP5 pre-amplifier.

As test cases, artificially damaged bearings are used. Damages of different sizes are introduced to the inner ring and outer ring (slots of 376 µm, 1880 µm each), and rolling elements (flattening by 94 µm and 376 µm) by wire erosion (see Fig. 3).

Fig. 3. Artificially damaged bearings, each w/o damage (left) and small damage (right); Inner ring damage, outer ring damage, and rolling element damage.

Measurements were conducted at two different speeds (800/2000 rpm) and with different damaged bearings mounted. In order to separate the effect of mounting a damaged bearing from the variability of the set up when demounting and mounting due to uncertainties in torque of screws, orientation of the damage at the bearing etc., measurements were repeated several times after demounting and mounting the bearing.

In total, data from 39 measurements was collected and used to train machine learning based classifiers as explained in the following Sections 5 and 6.

5. Data Preprocessing and Classifier Set Up

At first, one measurement per class was separated for validation. All measurements in the resulting validation dataset were measured at a rotation speed of 2000 rpm and the damaged bearings all had the smaller defect severity. The remaining 35 measurements with varying rotation speeds, failure severities and defect states were used to train the classifiers. This training dataset was again splitted into two parts: A stage 1 training dataset (5x no defect, 5x inner ring defect, 9x outer ring defect, 6x rolling element defect) and a stage 2 training dataset (1x no defect, 3x inner ring defect, 3x outer ring defect, 3x rolling element defect). An illustration of the division into datasets is shown in Fig. 4.

Measurements from both sensors were divided into snippets to create distinct samples for training and evaluation of the classifiers. An AE sample consists of 8000 consecutive values measured at a sampling rate of 390625 values per second (20.48 ms per sample), while a vibration sample consists of 4096 values, which corresponds to one second of measuring.

Vibration signals from the sensor mounted close to the bearing was used. For each vibration or AE sample, the mean and standard deviation was calculated and it was afterwards standardized using these values. From each standardized sample, the Fast Fourier Transformation (FFT) was calculated. For AE data, the frequency range between 512 Hz and 50000 Hz was extracted, and for the vibration data, the frequency range between 5 Hz and 512 Hz was used. The previously calculated mean and standard deviation values were appended to the FFT-transformed samples and thereby create a feature vector for each sample. An additional scaling of all datasets to a feature range between 0 and 1 was conducted based on the range parameters of the stage 1 training dataset. Examples of the measured sensor time series as well as their transformation into frequency domain are provided in Fig. 5.

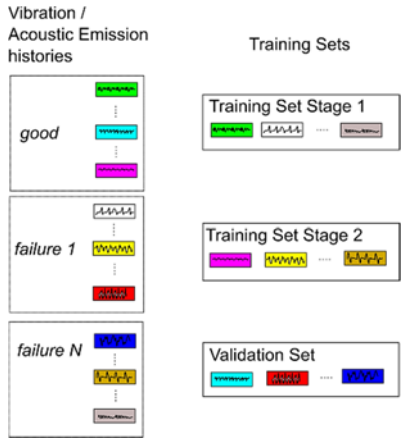


Fig. 1. Splitting of the measurements into a stage 1 and stage 2 training dataset as well as a validation dataset.

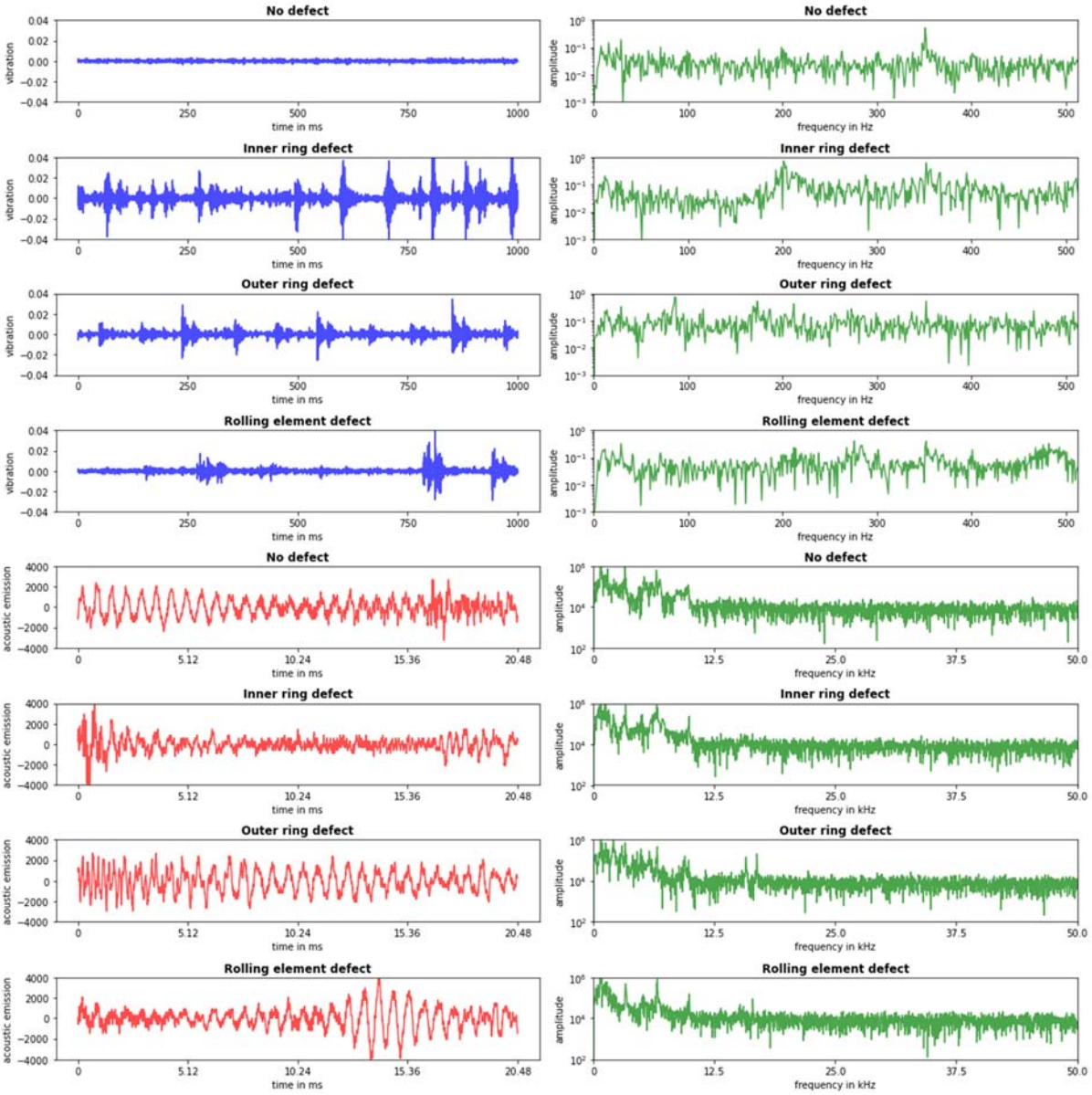


Fig. 2. Examples from the measurement data. Left column: Samples from all four used defect states in time domain. The upper four plots show vibration measurements, the lower four plots show AE data from the same measurement. Right column: FFT transformations of the samples in the left column.

All the sub classifiers mentioned in Section 3 are multi-layer perceptron. This type of neural network together with the frequency transformed input samples was chosen, since for vibration measurements it was shown that it can lead to classifiers with better generalization ability compared to e.g. convolutional neural networks which receive the time domain signals as input [2]. The AE classifier has 3 hidden layers, each with 1024 nodes, the vibration classifier has 2 hidden layers, also consisting 1024 nodes, respectively. Dropout was applied to all hidden layers of both classifiers to reduce overfitting. The combined classifier uses the activations of the last hidden layer of both the AE and the vibration classifier as one concatenated input vector. In addition, it has two hidden layers, each with 128 nodes. No dropout was applied to the combined classifier since enough variability was induced by the AE and vibration classifiers. During training phase of the combined classifier, the weights of the AE and the vibration classifier are frozen, to only fit the weights of the combined classifier.

6. Results of the Classifier Training

At first and as a baseline, both the AE classifier and the vibration classifier were trained on the full training dataset (including stage 1 and stage 2 training dataset). Each one got the scaled feature vectors from the respective sensor. Both classifiers achieve accuracies close to 100 % on the training dataset. The classification results of the trained classifiers on the unseen validation dataset are visualized in Fig. 6(a) and (b). While both classifiers correctly classify the defect free measurement and the measurement with the damaged rolling element, they both fail to classify the measurement of the outer ring defect. The inner ring defect can be recognized by the AE classifier, while the vibration classifier is only partially able to correctly classify this measurement.

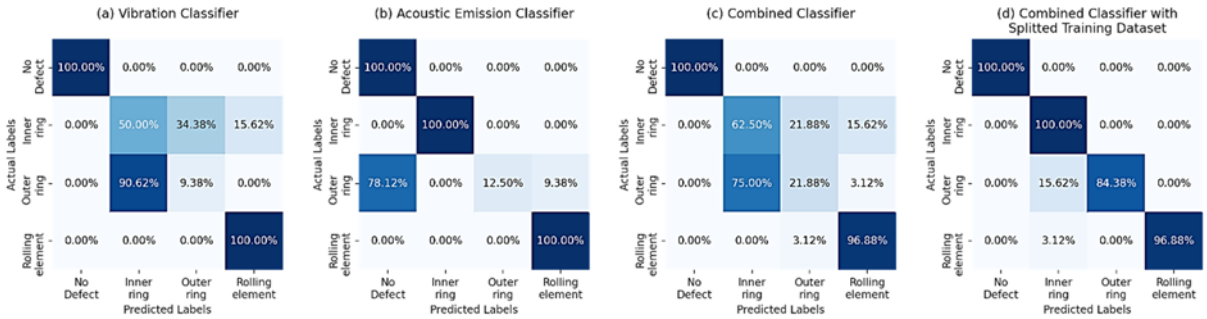


Fig. 3. Confusion matrices of the classification results on unseen validation measurements. (a) Vibration classifier, (b) AE classifier, (c) Combined classifier, where all sub classifiers were trained on the whole training dataset, (d) Combined classifier, where the AE- and the vibration sub classifiers were trained on one split of the training measurements and the combined classifier was trained on the other split of the training measurements.

Here we see a typical weak point of drive train CM systems: Since the frequency spectra of single samples inside one measurement differ only slightly, it is a quite easy task for any machine learning algorithm to recognize other samples from the same measurement. However, there are comparatively huge differences in frequency spectra between samples from different measurements even when they belong to the same defect class and especially, when the setup is dismantled and reassembled between the measurements as it was done here. It is therefore easier for the classifier to learn characteristic features of each measurement rather than to learn characteristic features of each defect state. A much larger number of measurements would be necessary to compensate for these effects and to achieve robust classifiers.

An approach to increase the robustness of the defect state classification is to combine the information from both used sensor modalities. This is done by the proposed combined classifier. In a first experiment the combined classifier is trained on the full training dataset. The pretrained vibration and AE classifiers thereby provide the input for the combined classifier. Their weights however remain unchanged during this procedure. The classification accuracies of the resulting classifier on the validation dataset is depicted in Fig. 6(c). Still, the inner ring defect and the outer ring defect cannot be classified correctly. And again, the reason for this result is the overfitting of the vibration and AE classifier. Since all classifiers receive input data from the same measurements, the combined classifier is not able to learn when to trust more the vibration classifier and in which cases the AE classifier is more reliable, because both of them are almost perfectly able to correctly classify samples from the dataset they are trained on.

To let the combined classifier learn when to trust which of the preclassifiers more, it needs to be trained with data from measurements unseen to the vibration and the AE classifier. For this purpose, the splitting of the training dataset in the stage 1 and the stage 2 training datasets as described in Section 5 was conducted. At first, the vibration classifier and the AE classifier were trained on data from the stage 1 training dataset. The combined classifier was afterwards fitted to data from the stage 2 training dataset. As with the

validation data, both preclassifiers were not able to perfectly classify the measurements from the stage 2 dataset and therefore the combined classifier could learn when to trust which one more or less. The validation result of the combined classifier trained with the described procedure is shown in Fig. 6(d). A remarkable improvement in the classification accuracy is apparent which indicates that the resulting classifier is much more robust to changing vibration behavior caused by a remounting of the setup compared to the classifier approaches discussed before.

7. Discussion and Outlook

Using the presented approach, bearing damages at drive trains can be detected early, precisely and robustly. While both sensor types enable a defect state classification separately, their data fusion further improves the fault detection sensitivity when they are trained on data from different measurements. In a next step, the concept needs to be evaluated on a broader and continuous rotation speed range as well as with differently sized drive trains. Additional sensor modalities (e.g. motor current, thermography) can increase the capabilities of the CM system further.

References

[1]. A. Dadouche, *et al.*, Sensitivity of air-coupled ultrasound and eddy current sensors to bearing fault detection, *Tribology Transactions*, Vol. 51, Issue 3, 2008, pp. 310-323.

[2]. O. Mey, W. Neudeck, A. Schneider, O. Enge-Rosenblatt, Machine learning-based unbalance detection of a rotating shaft using vibration data, in *Proceedings of the 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'20)*, Vienna, Austria, 08-11 Sept. 2020, pp. 1610-1617.

[3]. C. Li, *et al.,* Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals, *Mechanical Systems and Signal Processing*, Vol. 76, 2016, pp. 283-293.

[4]. R. B. Randall, J. Antoni, Rolling element bearing diagnostics – A tutorial, *Mechanical Systems and Signal Processing*, Vol. 25, Issue 2, 2011, pp. 485-520.

[5]. M. Kang, *et al.*, Reliable fault diagnosis for low-speed bearings using individually trained Support Vector Machines with kernel discriminative feature analysis, *IEEE Transactions on Power* Electronics, Vol. 30, Issue 5, 2015, pp. 2786-2797.

(012)

Experimental Validation of Petri Nets Based Regulation Control in a Small-scale Manufacturing System

J. M. Chavez [1], A. C. Gaona [1], C. R. Vázquez [1] and A. Ramírez-Treviño [2]
[1] Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias, Ramón Corona 2514, 45038, Zapopan, Mexico
[2] CINVESTAV, Unidad Guadalajara, Av. Del Bosque 1145, 45019, Zapopan, Mexico
E-mail: cr.vazquez@tec.mx

Summary: New industrial paradigms and technologies bring new possibilities for the automation of systems, increasing the flexibility of the production, customization of products, reducing set-up times among other value drivers. Nevertheless, these new possibilities imply complex challenges. For this, the use of formal methods can provide effective solutions to model, analyze and control large and complex manufacturing systems. In particular, we have proposed in previous works a control design approach based on Petri nets, called regulation control. In this work, we describe the application of this regulation control methodology for the modelling, specification and synthesis of a control program for a small scale manufacturing cell. The goal is to validate different algorithms and methodologies of this approach, and to detect omissions and opportunities to extend the regulation control theory.

Keywords: Automation, Manufacturing systems, Petri nets, Regulation control.

1. Introduction

Since the introduction of the Supervisory Control Theory (STC) in the '80s, many researchers have addressed the control problem of systems that evolve by the occurrence of events, named Discrete Event Systems (DES), such as automated manufacturing systems. The SCT was firstly proposed for systems modeled by finite state automata, where events are represented by arcs with labels, describing the possible transitions between the system's states. The synthesis of supervisors relies on regular languages analysis. In detail, a trajectory in the system is represented by a sequence of events occurrences, i.e., a sequences of labels associated to those events, such sequence is called a word. The set of all possible sequences in the system is called the system's language. In this control framework, the specification is defined as a set of safe event sequences, i.e., a specification language. Thus, the controller is an agent that prevents the occurrence of certain events, named controllable, in order to make the system to evolve inside the specification language, i.e., to maintain a safe behavior. The controller can be realized as a state automaton whose synchronization with the system leads to the required behavior. The SCT has been extended to consider several scenarios, e.g., distributed control, hierarchical control, control under partial observation, time notion, etc. A recent survey of SCT can be found in [1]. Nevertheless, there are some practical difficulties when applied the SCT in automation problems. First, the engineer has to be familiar with the SCT and regular language theory. Secondly, modelling with automata requires to enumerate all the possible states in the system. This number of states grows exponentially w.r.t. the number of components in the system. For instance, if we have a system with two actuators, each one described by two states (On, Off), then we have a total of four states (the combination of the states of the components). Thus for a system with 30 actuators, we will have an automaton model of 2^{30} states!

In order to deal with the state explosion problem, the SCT has been extended to Petri nets (PN). PN is formalism for modelling and analyzing DES with concurrency. In a PN model, there are two kinds of nodes, named transitions and places: events are presented by transition nodes, potential local (component) states are represented by place nodes. The current state of the components in the system is indicated by a dot, named token or mark. Thus, if we model a system with 30 actuators, each one having two states, we will require only 60 place nodes, 60 transition nodes and 30 tokens. Thus, the number of nodes in the PN model grows linearly w.r.t. the number of components in the system. Nevertheless, for the application of the SCT in PNs, the engineer still requires a deep knowledge on the system behavior and SCT theory.

Other control approaches have been proposed for PN [2]. A particularly interesting approach is the so called Generalized Mutual Exclusion Constraints (GMEC). In this, the goal is to avoid that certain components reach certain states simultaneously, representing problematic (unsafe) states. For this, new places, named monitors, are connected to the system, imposing new pre-conditions to avoid such problematic states. This technique has been widely extended in order to avoid deadlock states in certain classes of PN that represent manufacturing systems. Nevertheless, the problem of eliminating deadlock states is nontrivial, since the computation of potential deadlocks is a NP problem and the addition of

monitors may introduce new deadlock states. In any case, the GMEC approach is only applied to avoid reaching undesired states.

In this work, we consider the regulation control approach for PN [3-5], which has been proposed in order to enforce a system to follow particular trajectories that fulfill with high-level specifications provided by the engineer. In this approach, the engineer requires neither a deep knowledge on PN theory nor on the system's behavior. We claim that the regulation approach is suitable for developing the control programs in automated systems. However, this technique is still under development.

The contribution of this work is the experimental validation of the modelling, specification and synthesis control methodologies for the regulation control approach presented in [3-5], in order to detect practical issues and omissions. For this, we used a small-scale manufacturing cell depicted in Fig. 1, which dispenses wooden caps with three different shapes (square, circle, octagon) to a conveyor, collocate RFID tags to the caps and drill ten bores in the caps by using a CNC drilling machine.

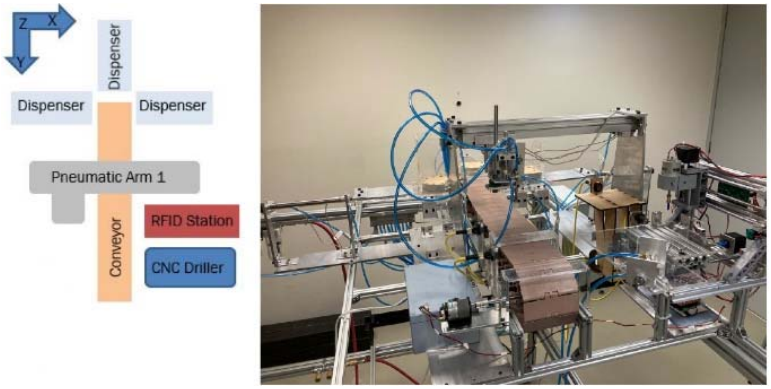


Fig. 1. Manufacturing cell.

This work is organized as follows. In Section 2, an introduction of the Petri net formalism is presented. Section 3 explains the existing regulation control methodologies for modelling, specifying and synthesizing controllers under this framework. Section 4 presents the experimental validation results, describing the omissions, inconsistencies and problems found. Finally, Section 5 provides some conclusions.

2. Petri Nets Definitions

The regulation control framework is based on *Interpreted Petri nets* models, which is an extension of PN that allows to represent input and output symbols in transitions and places, respectively. Let us provide some basic definitions.

Definition 1. A Petri net (*PN*) structure is a bipartite digraph represented by the 4-tuple $G = \langle P,T,I,O \rangle$, where $P = \{p_1,p_2,...,p_n\}$ is the finite set of places, $T = \{t_1,t_2,...,t_o\}$ is the finite set of transitions, $I: P \times T \rightarrow Z_{\geq 0}$ is a function representing the weighted arcs connecting places to transitions, $O: P \times T \rightarrow Z_{\geq 0}$ is a function representing the weighted arcs connecting transitions to places. $Z_{\geq 0}$ represents the non-negative integers.

Pictorially, places are represented by circles, transitions by rectangles, and arcs by arrows. The *incidence matrix* of a *PN* is a $|P| \times |T|$ matrix $\boldsymbol{C}$ defined such that $\boldsymbol{C}[i,j] = O(p_i,t_j) - I(p_i,t_j)$. A *PN* is said to be a *state machine* if $|{\bullet}t_j| = |t_j{\bullet}| = 1$, $\forall t_j \in T$.

Definition 2. Given a *PN* structure, the *marking* distribution is defined as a function $M: P \rightarrow Z_{\geq 0}$, where $M(p_i)$ represents the number of tokens residing inside the place p_i (tokens are depicted as dots). The marking distribution is expressed as a column vector $\boldsymbol{M}$ of length $|P|$, where $\boldsymbol{M}[i] = M(p_i)$, $\forall p_i \in P$. A *PN system* is the duple $\langle G,\boldsymbol{M}_0 \rangle$, where G is a *PN* structure and $\boldsymbol{M}_0$ the initial marking distribution. The marking distribution evolves according to the firing of transitions. A transition t_j is *enabled* at marking $\boldsymbol{M}_k$ if $\forall p_i \in {\bullet}t_j$, $\boldsymbol{M}_k(p_i) \geq I(p_i,t_j)$, this is denoted as $\boldsymbol{M}_k[t_j\rangle$. A transition t_j can fire if it is enabled. The firing of an enabled transition t_j leads to a new marking $\boldsymbol{M}_{k+1}$ that can be computed by

$$\boldsymbol{M}_{k+1} = \boldsymbol{M}_k + \boldsymbol{C}\boldsymbol{e}_j,$$

where $\boldsymbol{e}_j[i] = 0$ for $i \neq j$ and $\boldsymbol{e}_j[j] = 1$. This is denoted as $\boldsymbol{M}_k[t_j\rangle \boldsymbol{M}_{k+1}$.

Given a *PN* system $\langle G,\boldsymbol{M}_0 \rangle$, a sequence of transitions $\sigma = t_1t_2,...,t_k$ is said to be a firing sequence (or fireable) from $\boldsymbol{M}_0$ if there exist markings $\boldsymbol{M}_1, \boldsymbol{M}_2,..., \boldsymbol{M}_{k-1}$ such that $\boldsymbol{M}_0[t_1\rangle \boldsymbol{M}_1$, $\boldsymbol{M}_1[t_2\rangle \boldsymbol{M}_2,...,$ $\boldsymbol{M}_{k-1}[t_k\rangle$. The marking $\boldsymbol{M}'$ reached after the firing of a fireable σ from a marking $\boldsymbol{M}$ can be computed by the so-called *PN fundamental equation*

$$\boldsymbol{M}' = \boldsymbol{M} + \boldsymbol{C}\vec{\sigma},$$

where $\vec{\sigma}$ is a vector, named Parikh vector, defined as a column vector of size $|T|$ such that $\vec{\sigma}[j] = k$ if t_j is fired k times in the sequence σ. This is denoted as $\boldsymbol{M}[\sigma\rangle \boldsymbol{M}'$, moreover, $\boldsymbol{M}'$ is said to be *reachable* from $\boldsymbol{M}$.

A *PN* system is said to be *safe* if, for every reachable marking, there is at most one token in each place. A *PN* is said to be *deadlock-free* if every reachable marking enables at least one transition, otherwise it is said to have a deadlock.

Definition 3. An Interpreted Petri net (*IPN*) system is a 6-tuple $Q = \langle G,\mathbf{M}_0,\Sigma_I,\Sigma_O,\lambda,\varphi\rangle$, where $\langle G,\mathbf{M}_0\rangle$ is a *PN* system, and:

- Σ_I is the input alphabet, where each element of the set Σ_I is an input symbol;
- Σ_O is the output alphabet, where each element of the set Σ_O is an output symbol;
- $\lambda: T \rightarrow 2^{\Sigma_I}$ is the input-labeling function of transitions (a single transition can be associated with more than one symbol from the input alphabet Σ_I);
- $\varphi: P \rightarrow 2^{\Sigma_O}$ is the labeling function of places (a single place can be associated with more than one output symbol).

If $\lambda(t) \neq \emptyset$, the transition t is said to be *controllable*, otherwise it is *uncontrollable.*

The evolution of an *IPN* is similar to that of the *PN* system with the addition of the following rules:

1) a symbol $a \in \Sigma_I$ is said to be *indicated* if either it is activated by an external device (for instance a controller or a user) or $\exists p \in P$ such that $a \in \varphi(p)$ and $M(p)>0$ (notice that Σ_I may include symbols of Σ_O);
2) if $\lambda(t_j) = a$ is indicated and t_j is enabled then t_j *must* fire;
3) if $\lambda(t_j) = a$ and t_j is enabled, but the symbol $\lambda(t_j)$ is not indicated, then t_j cannot fire;
4) if $\lambda(t_j) = \emptyset$ and t_j is enabled, then t_j *can* fire at any moment;
5) at any reachable marking $\mathbf{M}_k$, an external observer reads the symbols $\varphi(p_i)$ if $M(p_i)>0$.

An *IPN* is said to be *event-detectable* if $\forall t_i,t_j \in T$ it holds either they are associated to different input symbols or their firings produce different marking changes.

Example 1. Fig. 2 shows an IPN Q where $P = \{p1, p2, p3, p4\}$, $T = \{t1, t2, t3, t4\}$, $\Sigma_I = \{a,b\}$, $\Sigma_O = \{A,B,C,b\}$. The incidence matrix of Q is given by

$$\boldsymbol{C} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

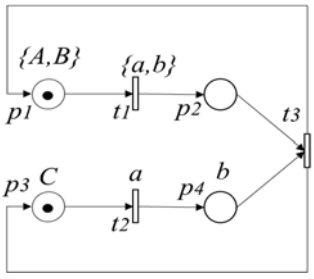


Fig. 2. *IPN* system described in Example 1.

The initial marking is $\mathbf{M}_0 = [1,0,1,0]^T$, describing that p_1 and p_3 have a token but p_2 and p_4 have not. Thus, at the initial marking, the *IPN* indicates the output signals $\{A,B,C\}$. Now, at the initial marking, both t_1 and t_2 are enabled by the marking, however, none of them can fire since both have input symbols that must be indicated (they are controllable). In particular, t_1 requires that both symbols a and b be indicated. If an external agent indicates the symbol a, only t_2 is fired. The firing of t_2 transfers a token from p_3 to p_4. The new marking can be computed as $\mathbf{M}_1 = \mathbf{M}_0 + \mathbf{C}\boldsymbol{e}_2 = [1,0,0,1]^T$, where $\boldsymbol{e}_2 = [0,1,0]^T$ (i.e., the elementary vector associated to t_2), thus p_1 and p_4 are the marked places. At this new marking, the symbols $\{A,B,b\}$ are indicated. Then, t_1 is fired if an external agent indicates the symbol a again (since b is indicated by the current marking). The firing of t_1 transfers a token from p_1 to p_2. At this new marking, t_3 is enabled since its input places have tokens and it is uncontrollable (it is not labelled), thus it can fire at any time, leading the system to the initial marking. In this way, the sequence $\sigma = t_2t_1t_3$ is a sequence fireable from the initial marking. As an example, $\sigma' = t_1t_2t_1$ is not fireable from the initial marking, since t_1 cannot fire twice without firing t_3.

3. Regulation Control Methodologies

In the regulation control framework, the system to be controlled, named the *Plant*, and the desired behavior, named the *Specification*, are both modelled as *IPN*. The goal is to compute a controller that drives the Plant in such a way that its output becomes equal to the Specification's output, i.e., the Plant's output tracks the Specification's output.

Definition 4. The *Plant* is a safe event-detectable *IPN* $Q = \langle G,\mathbf{M}_0,\Sigma_I,\Sigma_O,\lambda_I,\varphi\rangle$ that models the system to be controlled, where $G = \langle P,T,I,O\rangle$. It is assumed that two plant places cannot generate the same output symbol (i.e. a sensor cannot be turned on by two different variables), moreover, the input and output plant alphabets are disjoint (i.e., $\Sigma_I \cap \Sigma_O = \emptyset$).

In [3], we have introduced a modelling methodology for electro-pneumatic systems. In accordance, the manufacturing cell is modelled as the synchronous composition of the dispensers, pneumatic arm, CNC driller and RFID stations PN models. Neither the conveyor nor input/output buffers are included in the PN model. The model of each station is obtained by selecting the PN model for each of their pneumatic actuators from the library introduced in [3], shown in Figs. 3-5. The resulting model is depicted in Fig. 6.

Each of the three dispensers involves a double-effect pneumatic-actuator. The pneumatic arm consists of three double-effect pneumatic actuators (one for each axis X-Y-Z) and one suction cup driven by a vacuum ejector valve. The RFID station involves two pneumatic-actuators. The CNC driller includes one double-effect pneumatic-actuator to secure the caps to be drilled. Moreover, the CNC operation is represented by a Petri net with two local states, "home" and "drilling". The resulting PN model is denoted as Q^P.

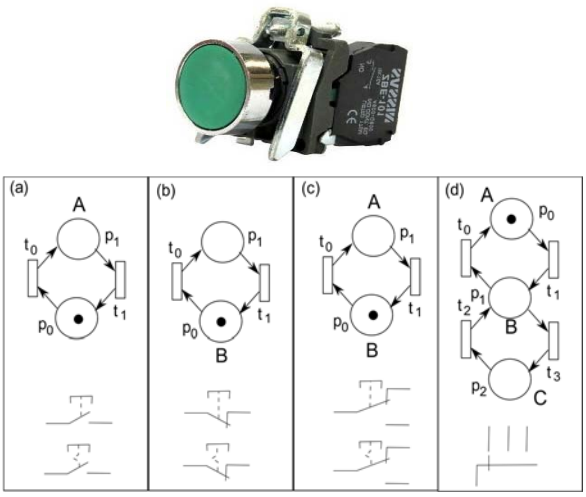


Fig. 3. *IPN* models for different kinds of switches.

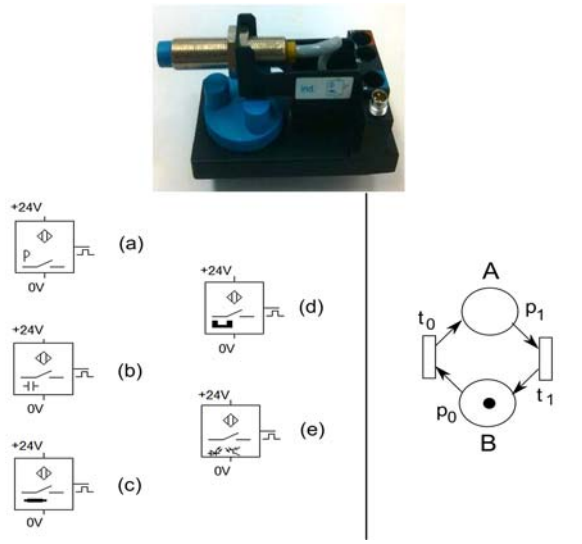


Fig. 4. Sensor symbols and their *IPN* model for a pressure sensor (a), capacitive sensor (b), inductive sensor (c), magnetic sensor (d), and optical sensor (e).

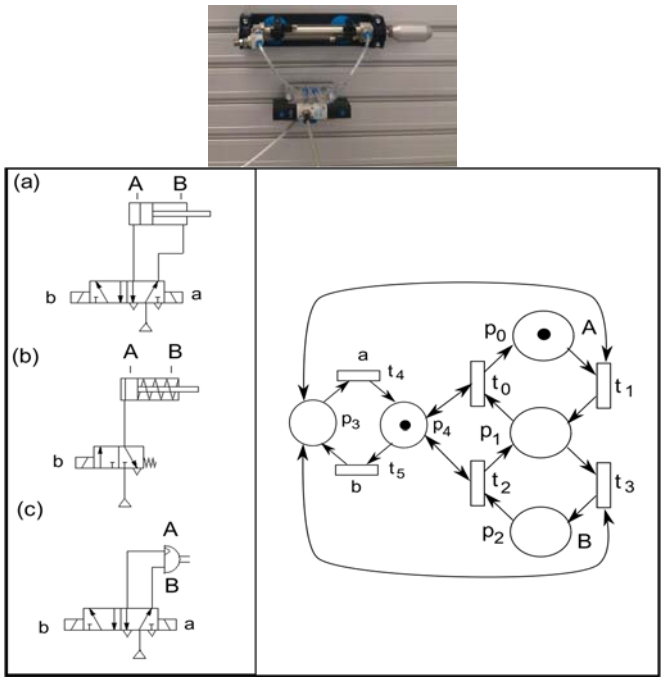


Fig. 5. *IPN* models for electro-pneumatic assemblies: double acting actuator controlled by a 5-ways 2-positions valve (a), spring return actuator controlled by a 3-ways 2-positions spring return valve (b), and a rotary actuator controlled by a 5-ways 2-positions valve (c).

Once the model of the Plant is defined, the next step is the specification definition, i.e., to describe the required behavior of the Plant.

Definition 5. Given a plant as defined in Definition 4, a *Specification* is a deadlock-free state-machine *IPN* $Q^S = \langle G^S, \mathbf{M}_0^S, \Sigma_I^S, \Sigma_O^S, \lambda^S, \varphi^S \rangle$. The output alphabet fulfills $\Sigma_O^S \cap \Sigma_I = \emptyset$. The specification input alphabet contains symbols from the plant's alphabets and other external symbols, whose set is denoted as $\Sigma_{ext} = \Sigma_I^S \setminus \{\Sigma_I \cup \Sigma_O\}$.

The specification is built by following the methodology of [5]. First, the cell is split in four workstations: dispensers, pneumatic arm, RFID and CNC driller. For each workstation, a set of tasks is defined. The dispensers perform three tasks: dispensing a rectangular cap (task named Rectangle), dispensing a circled cap (Circle), dispensing and octagon cap (Octagon). The CNC performs one task: drilling a cap (Drill). The pneumatic arm can perform four tasks: moving a cap from conveyor to RFID (In_2_RFID), moving a cap from RFID to CNC (RFID_2_CNC), moving a cap from conveyor to CNC (In_2_CNC) and moving a cap from CNC to conveyor (CNC_2_Out). Each task is described by a sequence of states that are reached after the task operations execution, each state is represented by a combination

of sensors activation (e.g., the task Rectangle is described by the sequence HomeDisp > APSQR.B > HomeDisp; where the HomeDisp state is characterized by the Reed sensors of the dispensers actuators detecting their retracted position, and the APSQR.B state is characterized by the Reed sensor of the square-cap dispenser at its extended position, as can be seen in the model of Fig. 6).

Finally, guards are defined for state transitions, representing proximity sensors and push-buttons that must be active in order to allow the execution of the operation. This information is captured in tables as shown in Fig. 7. Later, those tables are translated into PN models, one for each workstation. Fig. 8 shows the resulting IPN specification for the Cap's Dispensers workstation, in which each operation is represented by a place, and the three tasks lead to three branches from the HomeDisp place. The composition of those models leads to the PN specification denoted as Q^S.

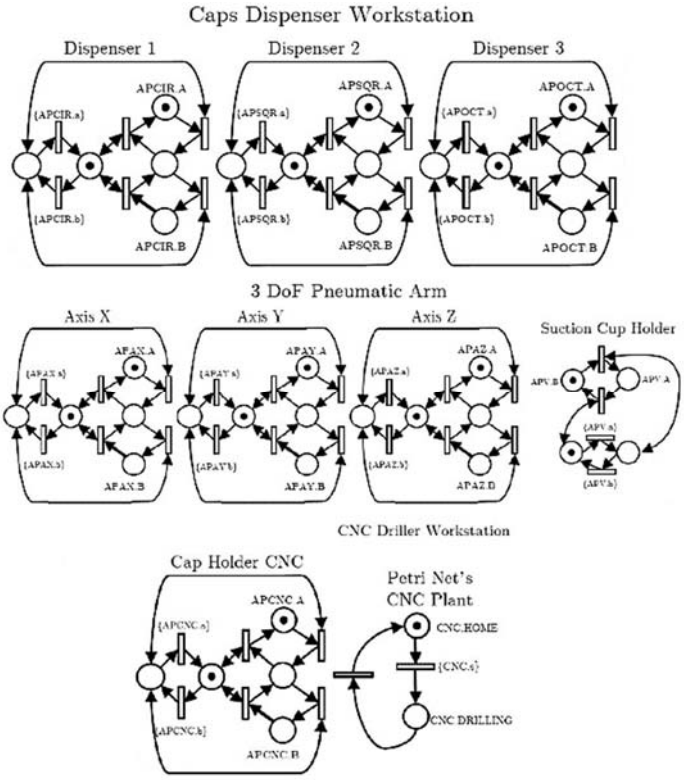


Fig. 6. *IPN* model for Plant, including the cap dispensers workstation, and 3 DoF pneumatic arm and the CNC driller workstation.

Workstation	Task name	Operation Sequence
Cap's Dispensers	Circle	HomeDisp, APCIR.B
Cap's Dispensers	Rectangle	HomeDisp, APSQR.B
Cap's Dispensers	Octagon	HomeDisp, APOCT.B
3 DoF Arm	In_2_RFID	HomeArm, take, APAX.B, APV.B
3 DoF Arm	In_2_CNC	HomeArm, take, APAX.B, APAY.B, place
3 DoF Arm	RFID_2_CNC	HomeArm, APAX.B, APV.A, APAY.B, place
3 DoF Arm	CNC_2_Out	HomeArm, APAX.B, APAY.B, take, APAX.A, place
CNC driller	Drill	HomeCNC, APCNC.B, CNC.FINISH

Task name	Event Conditions
Circle	(HomeDisp,APCIR.B, (CircleSignal.ON*SEMSTP.OFF))
Rectangle	(HomeDisp, APSQR.B, (SqrSignal.ON*SEMSTP.OFF))
Octagon	(HomeDisp, APOCT.B, (OctagonSignal.ON*SEMSTP.OFF))
In_2_RFID	(HomeArm, take,(SMP.OFF*SPRFID.ON*TagRFID.OFF*SEMSTP.OFF))
In_2_CNC	(HomeArm, take,(SMP.OFF*SPRFID.ON*TagRFID.ON*CNC.HOME*SPCNC.ON*SEMSTP.OFF))
RFID_2_CNC	(HomeArm, APAX.B,(CNC_Finish_Signal*CNC.HOME*SPCNC.ON*SEMSTP.OFF))
CNC_2_Out	(HomeArm, APAX.B,(CNC_Finish_Signal*SPRFID.ON*SPCNC.OFF*SPF.ON*SEMSTP.OFF))
Drill	(HomeCNC, APCNC.B,(CNC_Begin_Signal*SPCNC.OFF*SEMSTP.OFF))

Subtask name	Operation Sequence	Event Conditions
take	APAZ.B, APV.A, APAZ.A	-
place	APAZ.B, APV.B, APAZ.A	-

Fig. 7. Tables for defining the specification. Shortcuts are: HomeDisp = {APCIR.A,APSQR.A,APOCT.A}, HomeArm = {APAX.A, APAY.A,APAZ.A, APV.A}, HomeCNC = {APCNC.A,CNC.HOME}.

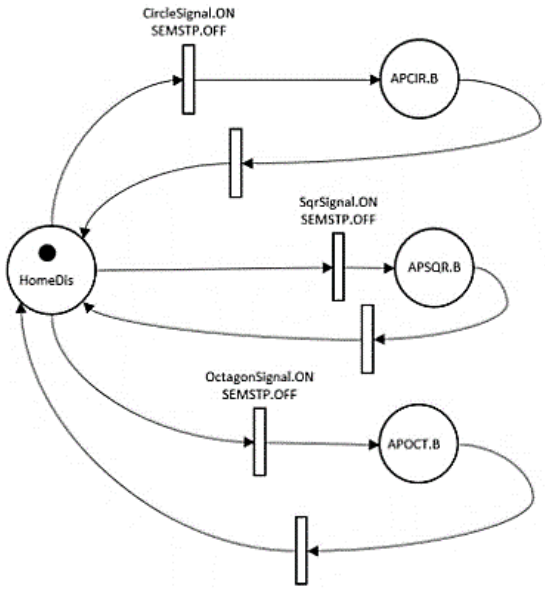


Fig. 8. IPN specification for the Cap's Dispensers workstation, obtained from the data of Fig. 7.

In [4], an algorithm for the automatic synthesis of a regulator controller is introduced. The application of such algorithm to the PN model Q^P and the obtained specification model Q^S results in another PN that represents the controller, denoted as Q^C. The synchronous composition of the plant, specification and controller represents the closed-loop system, which can be simulated to verify the correct operation. On the other hand, the synchronous composition of the specification and the controller represents the control program, which must be translated to a PLC code for implementation. Fig. 9 shows the synchronous composition of the specification and controller for the Cap's Dispensers workstation. Notice that the new places (controller places) include labels associated to controllable transitions, i.e., these places command the activation of actuators.

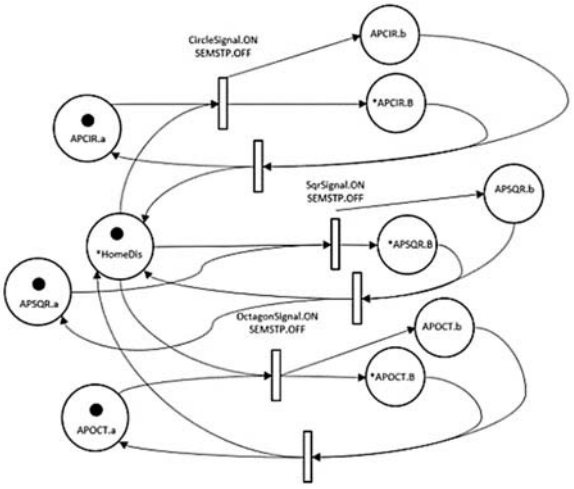


Fig. 9. IPN representing the synchronization of the specification and the controller for the Cap's Dispensers workstation. Labels with symbol * should be eliminated, here are kept for relating the specification places.

4. Validation

The modelling methodology was firstly analyzed. For this, we performed a classical model validation as commonly achieve for continuous control systems, i.e., we manually activate combinations of actuators (input signal) in the system and simulate the same input into the PN model; then we compare the evolution of the physical plant with that of the simulation. The results are summarized in the Table 1.

Table 1. Results after model validation.

Workstation	Discrepancies	Omissions	Dead-locks
Dispensers	-	×	–
Arm	×	×	–
Driller	–	×	–
Complete Cell	×	×	–

For instance, the discrepancy at the arm model is due to the suction cup with its ejector valve, whose behavior does not correspond to the original PN model, since the valve has an internal control-loop for saving energy. An important omission found in the dispensers model is due to that two or more dispenser actuators cannot operate simultaneously, since they dispatch a cap on the same physical place on the conveyor, but the model allows the simultaneous operation of two or more dispensers since neither the conveyor nor output buffers are modelled.

The next step was to analyze the specification mythology [5]. For this case, we obtained a set of four disconnected state machine PN as specification. For the synthesis of the controller, it would be convenient to transform those state machines into an equivalent single one. However, we found that even for this small Plant, such approach will lead to an inconveniently large state machine (348 places). Thus, we decided to maintain the specification as four independent PN.

In order to check the controllability of the tasks, we manually execute each one, i.e., by manually activating each actuator. This allowed us to detect a controllability issue in the CNC specification. In detail, the CNC specification requires to reach a state when the CNC is drilling, however it is not possible to maintain such state indefinitely, as required by the synthesis algorithm, since the CNC will immediately return to its home state after finishing the drilling of all the bores. This issue was avoided by adding a virtual binary-state variable to the CNC workstation model in order to detect when the CNC has visited its drilling state, and modifying the specification by considering this signal instead of the CNC drilling signal.

Finally, we applied the control synthesis algorithm of [4] for each workstation, obtaining thus four controllers. Next, we translated the control programs to ladder language. Finally, we implemented the four controllers in a single PLC device, obtaining thus a

distributed control architecture without communication [4]. In this case, we did not found any issue in the operation of the controlled system. However, we realized that the distributed control scheme worked as expected due to the fact that the conveyor synchronizes the workstations, and that the workstations do not compete by common resources. In other case, a coordinated control would be required to ensure a process specification [4].

5. Conclusions

Here, we reported the application of modelling, specification and control synthesis methodologies to a manufacturing cell for enforcing regulation control based on Petri nets. We found that the modelling and specification methodologies must be extended in order to consider buffers and conveyors, and to deal with components that are not controllable. Moreover, we found that a centralized control approach, in which the specification can be defined as a single state machine PN, is unsuitable even for small cases. Thus, the development of decentralized control synthesis techniques is of paramount importance for the applicability of the regulation control theory in real cases. After making some minor adjustments on the methodologies, we were able to synthesize a controller that achieves the control goal, validating thus the regulation control methodology in a general way.

Acknowledgements

The research leading to these results has received support from the Conacyt Fondo Sectorial de Investigación para la Educación, project number 288470.

References

[1]. W. M. Wonham, K. Cai, K. Rudie, Supervisory control of discrete-event systems: A brief history, *Annual Reviews in Control*, Vol. 45, 2018, pp. 250-256.

[2]. A. Giua, M. Silva, Petri nets and automatic control: A historical perspective, *Annual Reviews in Control*, Vol. 45, 2018, pp. 223-239.

[3]. C. Vázquez, J. Gómez-Castellanos, A. Ramírez-Treviño, Petri nets tracking control for electro-pneumatic systems automation, *Lecture Notes in Electrical Engineering*, Vol. 613, 2019, pp. 503-525.

[4]. D. Guevara-Lozano, C. Vázquez, A. Ramírez-Treviño, Towards decentralized control in Petri nets, in *Proceedings of the 24th International Conference on Emerging Technologies and Factory Automation (ETFA'19)*, Zaragoza, Spain, Sept. 2019, pp. 428-435.

[5]. D. Guevara-Lozano, C. Vázquez, A. Ramírez-Treviño, Automatic specification generation for tracking control in interpreted Petri nets, in *Proceedings of the 7th International Conference on Control, Decision and Information Technologies (CoDIT'20)*, Prague, Czech Republic, July 2020, pp. 341-346.

(013)

Measurements of Geometric Characteristics on Machine Tool as an Element of Closed Door Technology

P. Gierlak [1], A. Burghardt [1], K. Kurc [1], M. Muszyńska [1], D. Szybicki [1] and G. Bomba [2]
[1] Rzeszow University of Technology, Faculty of Mechanical Engineering and Aeronautics,
al. Powstańców Warszawy 8, 35-959 Rzeszów, Poland
[2] Pratt & Whitney Rzeszów S.A., ul. Hetmańska 120, 35-078, Rzeszów, Poland
Tel.: + 48 17 865 18 54
E-mail: pgierlak@prz.edu.pl

Summary: The article concerns the implementation of the closed door technology concept in the production of thin-walled aircraft transmission housings. High requirements regarding dimensions and shape impose the necessity to minimize the influence of the human factor in the technological process of manufacturing these parts. The article discusses the difference between the standard housing machining process and the closed door process, with particular emphasis on the specificity of manufactured parts. The concept of the implementation of the procedure for measuring geometric characteristics on a machine tool supported by a neural-fuzzy system is presented. The purpose of the neural-fuzzy system is to correct the results of measurements carried out on the machine tool in such a way that they are as close as possible to the results obtained with the use of the coordinate measuring machine. This makes it possible to reduce the number of measurements performed on the measuring machine and thus eliminate the bottleneck problem in the process. The results of the application of the neural-fuzzy system for data from the actual production process are presented, which confirms the usefulness of the developed solution.

Keywords: Closed door technology, Aircraft accessory gearboxes, Geometric measurement, Neuro-fuzzy system.

1. Introduction

The concept of machining in accordance with closed door technology (CDT) is based on the best use of a production line built of autonomous machines, capable of continuous production with minimal human intervention. In the case of the manufacturing of products based on machining, such a line generally consists of a group of numerically controlled (CNC) machining centers, equipped with a whole range of devices and systems ensuring the autonomy of the system. These devices include the multi-pallet system integrating several centers into one whole, as well as the master system controlling the operation of the line, pallet system and CNC machining centers. Machine tools should be equipped with systems monitoring the condition of tools, e.g. by optical measurement with a laser beam or contact elements.

The standard equipment for CNC machining centers has become a set of measuring probes with dedicated software for their operation. This makes it possible to take measurements and record the results in the form of CNC system variables or measurement reports.

The software of multi-axis CNC machining centers is extended with software packages using, among others measurement of the position of the work piece on the machine and software adjustment of the machining system of the work piece to the system used in the CAM environment [6].

CDT is characterized by full use of production possibilities thanks to better planning of workshop tasks and the maximum limitation of the impact of the human factor on the technological production process. When machining *behind closed doors*, operators play a different role than in traditional machining approaches. They move up the skill chain to become machining cell coordinators, capable of operating several machines simultaneously and intended to keep the production line up and running through process activities. They focus on supervisory and control tasks with high added value, while the machines operate independently [1].

The following sections of the article describe the possibility of using CDT in the production of housings for aircraft accessory gearboxes (AGB), part of the accessory drive train. Section 2 describes the process of producing the housings, Section 3 presents the possibility of automatically measuring the housings on a machine tool, and Section 4 presents a computational model for improving the quality of the measurement process.

2. Manufacture of Aircraft Transmission Housings

Aircraft AGB housings (Fig. 1) are intended for the reciprocal arrangement of various elements, assemblies and mechanisms. In order to reduce the weight, the external shapes of the housings do not greatly differ from the outlines of the elements located inside. This results in complexity of the shape and specific technological requirements. In order to minimize the weight, the housings are made of thin-walled light alloys. During the machining process, they are subject to geometric deformations resulting, among others, from the release of stresses occurring in

them resulting from the casting process. The machining process is the second cause of stress formation [5].

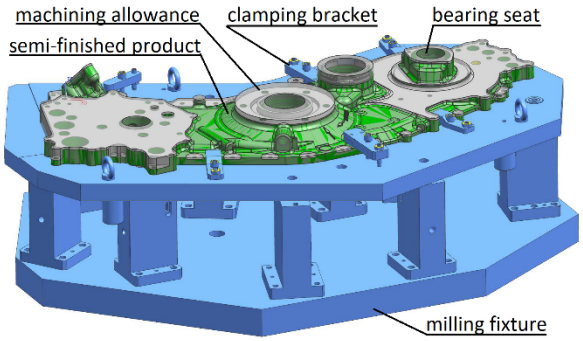


Fig. 1. AGB housing mounted in a milling fixture.

One of the most important geometric parameters of the transmissions is the arrangement of the bearing seats. It has a direct impact on the interaction of the gears, which is the main factor determining the durability of the transmission.

Machining of gear bearing seats is carried out with the use of precise machining centers. The position of the bearing seats is largely determined by the kinematic accuracy of the machine tool. The second factor influencing the position of the bearing seats is the deformation of the housing resulting from changes in the stress state during machining and after unclamping. Assuming that the machine tool used is allowed to carry out the process only when its kinematic accuracy is appropriate [1], the main emphasis should be placed on taking into account the influence of changes in the stress state, which result in an increase in the curvature of the transmission housing division surface.

Fig. 2 shows the standard process of machining AGB housings compared to the process carried out in CDT. Currently, the most difficult stage of the process from the point of view of implementing CDT is measurement operations, i.e. the last stage of the process in CDT. In principle, they should only take place on the machine tool. This requires replacing the coordinate measuring machine (CMM) with a machine tool equipped with measuring probes.

3. Determinants of Automatic Measurements

Precision CNC machining centers and CMMs have to some extent similar properties. They include, among others:

- Kinematics, because they are machines with mainly so-called Cartesian configuration;
- Precise measurement systems enabling the description of the position of a characteristic point in Cartesian space;
- Measuring equipment, i.e. measuring probes;
- Possibility of programming measurement cycles;
- The ability to generate reports directly or via system variables.

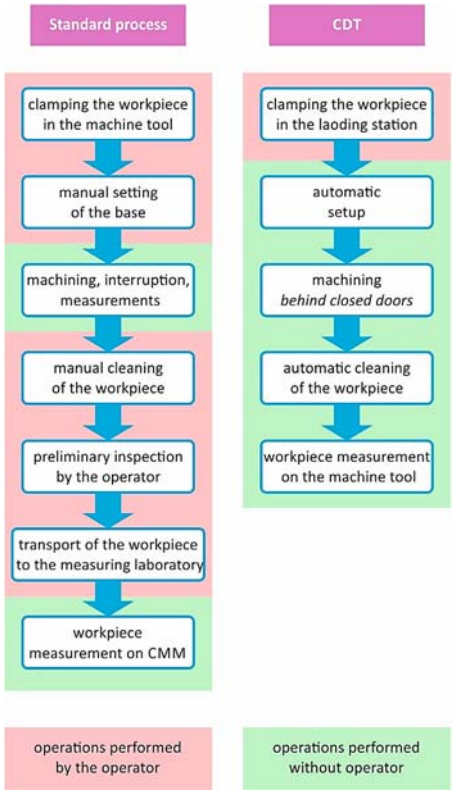


Fig. 2. Standard manufacturing process by machining and by CDT.

A significant challenge is the measurement of dimensional and shape characteristics on the machine tool and the generation of a measurement report that will ensure its identity with the measurement report generated on the basis of data collected with the CMM. An obstacle to such a task, however, is numerous factors that significantly complicate the mathematical and physical description of the condition of the AGB housing mounted on the machine tool after the machining operation. Of many, at least two basic factors can be mentioned. First, according to the design requirements, the housing should be measured in a loose state on a CMM. On the other hand, on the machine tool, the AGB housing is mounted in the machining device and is subject to clamping forces resulting from the pressure of the clamping elements. Second, the effect of the stresses released by the removal of the allowance and the stresses introduced during machining mean that the shape of the housing is different in the clamped state than in a loose state. Summing up, it is not possible to prepare a correct measurement report for an AGB based only on data collected on a machine tool with measuring probes. It should be emphasized that this is a specific problem in the machining of thin-walled components which are subject to high deformation, such as AGB housings.

Automation of measurements involving their inclusion in a CDT requires additional information about stresses. It is necessary to determine the relationship between the measurement results on a machine tool with the measurement results that would be obtained on a CMM in an unmounted state. Of course, direct information about stresses is impossible

to obtain, but the influence of stresses can be determined indirectly. First, it should be realized that changing the stresses in the machined housing changes the stresses in the components holding it. Additionally, after unclamping the housing, it deforms as a result of changes in stresses. As a result, the measurement results of the housing on the machine tool in the attached state and on the CMM in the unattached state differ. From these observations, there are two ways to obtain information, indirectly, about the change of stress in the housing. The first is the use of measuring systems in the housing mounting element. This makes it possible to determine changes in the housing mounting forces as a result of changes in stresses. This method requires significant modifications to the attachment for mounting the housing on the machine tool. The second method is based on the fact that the housing deforms after unclamping and changing the stresses. Both information on the change of the housing mounting forces and information on the deformation of the housing after unclamping can be used to build computational models allowing for the correction of geometric characteristics from the machine tool in a mounted state so that they correspond to the geometry of the housing in a free state.

The first method based on measuring the change in clamping forces still requires extensive research and is the target solution for the full implementation of CDT. The second method, described in this article, is a temporary solution, because its implementation means that not all of the measurement process is performed in CDT. This method consists in introducing a calibration measurement for each housing and using a computational model that takes into account both the calibration measurement and other measurements from the machine tool. The calibration measurement performed in an unattached state on the CMM consists in determining the curvature of the plane of the AGB housing division. On the basis of the authors' research [2] it was determined that taking into account the curvature of the plane allows us to find a correlation between the results of measurements on the machine tool and the CMM. The advantage of this solution is that several dozen geometric parameters are determined on the machine tool and only the curvature of the plane is determined by the CMM. This contributes to the elimination of the bottleneck in the manufacturing process: the CMM, which is used to control production on many processing lines [2].

4. Computational Model Supporting the Monitoring of Geometric Characteristics

Taking into account the characteristics of AGB housings, selected properties of machine tools and measuring machines, as well as the possibility of learning and interpreting the computational model, the use of an Adaptive Neuro-Fuzzy Inference System (ANFIS) was proposed [3]. The following assumptions were made for such a system:

- A separate ANFIS will be built to define each geometric characteristic in order to avoid large dimensionality of the problem as well as to obtain the possibility of model interpretation;
- The input data to ANFIS will be the measurement data obtained on the machine tool in an unmounted state and the curvature of the transmission housing division surface determined on the CMM in an unmounted state;
- Measurement data will be obtained on a machine tool according to measurement programs that are possibly convergent with measurement programs on CMM;
- The measuring probes used on the machine tool and CMM will have the same or similar parameters.

ANFIS combine the advantages of artificial neural networks and fuzzy systems [4]. It is possible to train such systems in the mappings on the basis of a data set and at the same time obtain a clear representation of knowledge in the form of a base of inference rules. In such a system, the subject of training are the parameters of premises and conclusions. Due to the generalizing properties of ANFIS, it is possible to build an inference model based on a relatively small set of experimental training data.

An example of the application of the discussed solution for determining the compliance of the position of the bearing seat with the base on the housing is presented below. The deviation of the seat position should not exceed 0.25 mm, which results from the technical requirements. Fig. 3 shows the position deviations of the bearing seats determined on the CMM (HPD_{CMM}) and the machine tool (HPD_{MT}). It is clear that the results vary considerably. What's worse, the results obtained from the measurements on the machine tool are in most cases "too optimistic" in relation to the actual state. After applying the learned ANFIS, the input of which is the deviation determined in the measurements on the machine tool (HPD_{MT}) and the curvature of the housing division surface (SC_{CMM}) (Fig. 5), deviations (HPD_{ANFIS}) were obtained with the values shown in Fig. 4. Additionally, the deviations were presented as a reference specified on the CMM.

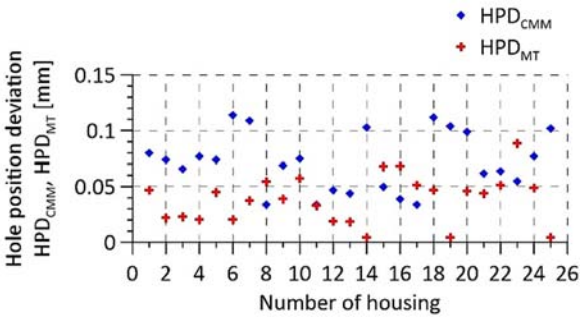


Fig. 3. Hole position deviations determined using the CMM and a machine tool.

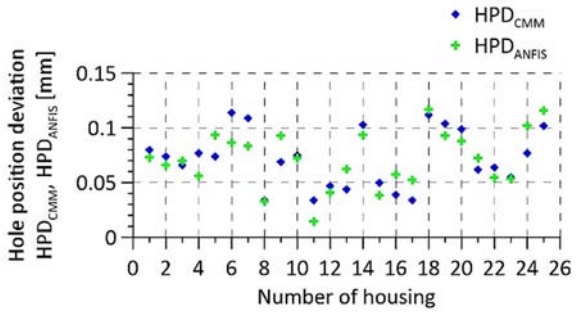


Fig. 4. Hole position deviations determined using CMM and ANFIS.

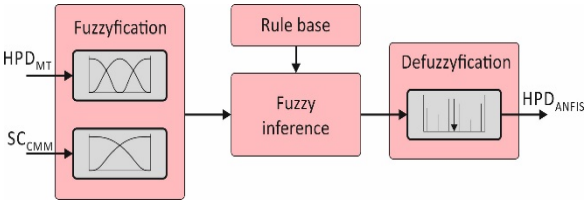


Fig. 5. ANFIS scheme.

Fig. 4 shows that the deviations determined by ANFIS are much closer to the results obtained with the CMM. The maximum difference between ANFIS results and CMM data is 0.0275 mm, while the maximum difference between machine results and CMM data is 0.0996 mm. Therefore, the use of ANFIS to correct the measurement process on the machine tool allowed us to obtain over three and a half times more accurate results.

5. Conclusions

The presented solution is a step towards the implementation of CDT in the production of thin-walled aircraft transmission casings. At the present stage, the solution enables the improvement of the measurement procedure. If the number of measurement characteristics determined by means of the CMM is much smaller than all the necessary characteristics, then the proposed solution eliminates the bottleneck problem and is rational from an economic point of view.

Further research is required to fully implement CDT. It will consist in the design, construction and implementation of a housing clamping device that will be equipped with a clamping force measurement system. Data from this system will be used to train ANFIS.

References

[1]. G. Bomba, P. Gierlak, Assessment of geometric accuracy of a 5-axis CNC machine in the context of machining aircraft transmission housings, *Universal Journal of Mechanical Engineering*, Vol. 8, Issue 5, 2020, pp. 258-264.

[2]. G. Bomba, P. Gierlak, dimensional control of aircraft transmission bodies using CNC machines and neuro-fuzzy systems, *Applied Sciences*, Vol. 9, Issue 19, 2019, 4094.

[3]. J. S. Jang, ANFIS: -network-based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23 No. 3, 1993, pp. 665-685.

[4]. H. X. Li, C. P. Chen, The equivalence between fuzzy logic systems and feedforward neural networks, *IEEE Transactions on Neural Networks*, Vol. 11, Issue 2, 2000, pp. 356-365.

[5]. A. Matras, M. Plaza, The FEM simulation of the thin walled aircraft engine corpus deformation during milling, *Proceedings of SPIE*, Vol. 10031, 2016, 100310B.

[6]. H. Wang, M. X. Zhou, W. Z. Zheng, Z. B. Shi, H. W. Li, 3D machining allowance analysis method for the large thin-walled aerospace component, *International Journal of Precision Engineering and Manufacturing*, Vol. 18, 2017, pp. 399-406.

(015)

Asymptotic Random Distortion Testing for Anomaly Detection

Dominique Pastor and Guillaume Ansel
IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238, France
Tel.: +33 229 00 11 11
E-mails: dominique.pastor@imt-atlantique.fr, guillaume.ansel@imt-atlantique.fr

Summary: In connection with cybersecurity issues in ICS, we consider the problem of detecting yet unknown attacks by presenting a theoretical framework for the detection of anomalies when the observations have unknown distributions. We illustrate the relevance of this framework with experimental results.

Keywords: Anomaly detection, Industrial control systems (ICS), Cybersecurity, Industry 4.0, Statistical hypothesis testing, Random distortion testing.

1. Introduction

Vulnerability of systems in Industry 4.0 and Smart Factories increases as the number of new threats grows with the number of connected devices, especially in ICSs (Industrial Control Systems). It is crucial to devise methods capable of reliably detecting novel types of attacks. Anomaly detection [1, 2] addresses this issue by considering possibly novel attacks as anomalies with respect to nominal system behaviors. However, the diversity of processes and the various variable users' habits and behaviors entail many deviations with respect to nominal system behaviors, even in the absence of attacks. Because of this variability, anomaly detection may yield too many false alarms. It is thus desirable to cast the anomaly detection problem into a theoretical framework to deal with deviations around a nominal model, with guaranteed performance and even optimality.

Although statistical hypothesis testing provides a "*statistically justifiable solution for anomaly detection*" even "*in an unsupervised setting without any need for labeled training data*" [1], they however assume that observations obey specific distributions, which is questionable in practice because of the aforementioned variations around nominal models, even in absence of attacks.

The Random Distortion Testing (RDT) [3] aims to overcome the aforementioned limitations of statistical methods and anomaly detection. This framework incorporates the existence of unknown deviations and deals with fully unknown probability distributions for the observation. The observation is assumed to result from some signal with unknown distribution observed in additive and independent Gaussian noise, and the RDT approach is optimal with guaranteed performance to decide whether the noisy signal drifts by too much from a certain deterministic model. Since RDT assumes a perfectly known noise distribution, we present below theoretical and experimental results to upgrade the original RDT framework to the White Gaussian Noise (WGN) case with estimated standard deviation.

Notation and terminology. $M(\Omega, \mathbb{R}^d)$ denotes the set of all d-dimensional real random vectors defined on the probability space $(\Omega, \Sigma, \mathbb{P})$. $Q_{d/2}$ is the generalized Marcum function. $\forall(\theta_0, \rho) \in \mathbb{R}^d \times (0, \infty), S_\rho = \{y \in \mathbb{R}^d : \|y - \theta_0\|_2 = \rho\}, B_\rho = \{y \in \mathbb{R}^d : \|y - \theta_0\|_2 \leq \rho\}$ and $B_\rho^c = \{y \in \mathbb{R}^d : \|y - \theta_0\|_2 > \rho\}$. For all $t > 0$, we define the test $T_t : \mathbb{R}^d \to \{0, 1\}$ by:

$$\forall y \in \mathbb{R}^d, T_t(y) = \begin{cases} 1 \text{ if } \|y - \theta_0\|_2 > t \\ 0 \text{ otherwise} \end{cases}$$

2. The RDT Approach in the WGN Case

The RDT problem can be stated as follows [3]:

$$\begin{cases} \textbf{Data model: } \exists (Y, \Theta) \in M(\Omega, \mathbb{R}^d)^2, \exists X \sim N(0, \sigma^2 \boldsymbol{I}_d), \\ \quad \begin{cases} \Theta \text{ and } X \text{ are independent,} \\ Y = \Theta + X, \\ \forall y \in \mathbb{R}^d, \exists \omega \in \Omega, y = Y(\omega) \end{cases} (1) \\ \textbf{Testing problem: } \text{Given } y = Y(\omega), \text{ test:} \\ \qquad H_0 : \Theta(\omega) \in B_\tau \text{ vs. } H_1 : \Theta(\omega) \in B_\tau^c \\ \text{with } \tau > 0 \text{ and } \theta_0 \in \mathbb{R}^d \end{cases} \quad (1)$$

The existence of optimal tests for the RDT problem is established via the notions of *conditional size* and *conditional power* defined as follows.

Definition 1. *For* $\rho > 0$ *and* $T : \mathbb{R}^d \to \{0, 1\}$, *we set:*
Conditional power:

$$\forall \Theta \in M(\Omega, \mathbb{R}^d), \mathbb{P}\big[T(\Theta + X) = 1 \mid \Theta \in S_\rho\big],$$

Conditional size:

$$\alpha_T = \sup_{\Theta \in M(\Omega, \mathbb{R}^d) : P[\Theta \in B_\tau] \neq 0} \mathbb{P}[T(\Theta + X) = 1 \mid \Theta \in B_\tau]$$

Definition 2. *Given* $\Theta \in M(\Omega, \mathbb{R}^d)$ *and* $\rho \geq 0$, *a test* T *is said to have constant conditional power function* (CCPf) *given* $\Theta \in S_\rho$ *if for every* $\theta \in S_\rho$:

$$\mathbb{P}\left[T(\Theta + X) = 1 \mid \Theta \in S_\rho\right] = \mathbb{P}[T(\theta + X) = 1]$$

We can then exhibit optimal tests as follows.

Theorem 1. *Given* $\gamma \in (0,1)$, *if* $\lambda_\gamma(\tau)$ *is such that* $Q_{d/2}\left(\tau, \lambda_\gamma(\tau)\right) = 1 - \gamma$, $T_{\lambda_\gamma(\tau)}$ *is optimal in that:*

i. $\alpha_{T_{\lambda_\gamma(\tau)}} = \gamma$;

ii. $\forall \Theta \in M(\Omega, \mathbb{R}^d)$ *and for* $\mathbb{P}(\|\Theta - \theta_0\|_2)^{-1}$ - *almost every* $\rho > \tau$, $T_{\lambda_\gamma(\tau)}$ *has CCPf given* $\Theta \in S_\rho$ *and for all* T *with* $\alpha_T \leq \gamma$ *and CCPf given* $\Theta \in S_\rho$:

$$\mathbb{P}\left[T_{\lambda_\gamma(\tau)}(\Theta + X) = 1 \mid \Theta \in S_\rho\right] \geq \geq \mathbb{P}\left[T(\Theta + X) = 1 \mid \Theta \in S_\rho\right]$$

3. Asymptotic RDT

For any $t > 0$, let $\tilde{T}_t$ be the function defined by:

$$\begin{array}{rcl} \tilde{T}_t : \mathbb{R}^d \times \mathbb{R}^d \times (0.\infty) & \to & \{0,1\} \\ (y, \theta, \sigma) & \mapsto & \begin{cases} 1 \text{ if } \|y - \theta\|_2 > \sigma t \\ 0 \text{ otherwise} \end{cases} \end{array}$$

We hereafter consider consistent estimators $\hat{\theta}_n$ and $\hat{\sigma}_n$ of θ_0 and σ_0 respectively. Set $Z = (Y, \theta_0, \sigma_0)$ and $\forall n \in \mathbb{N}, Z_n = (Y, \hat{\theta}_n, \hat{\sigma}_n)$. $\forall \rho \geq 0$ and $\forall n \in \mathbb{N}$, define $\Pi_n(\rho, \cdot)$ and $\Pi(\rho, \cdot)$ by setting for all Borel set $A \subset \mathbb{R}^{2d+1}$:

$$\begin{array}{rcl} \Pi_n(\rho, A) & = & \mathbb{P}[\, Z_n \in A \mid \|\Theta - \theta_0\|_2 = \rho\,] \\ \Pi(\rho, A) & = & \mathbb{P}[\, Z \in A \mid \|\Theta - \theta_0\|_2 = \rho\,] \end{array}$$

Theorem 2 (Asymptotic level)**.** If $S = \{\Xi \in M(\Omega, \mathbb{R}^d) : \forall n \in \mathbb{N}, \Xi$ and $(\hat{\theta}_N, \hat{\sigma}_N)$ *are independent*$\}$, *then*:

$$\limsup_n \sup_{\Xi \in S : \mathbb{P}[\Xi \in B_\tau] \neq 0} \mathbb{P}\left[\tilde{T}_{\lambda_\gamma(\tau)}(\Xi + X, \hat{\theta}_N, \hat{\sigma}_N) = 1 \mid \Xi \in B_\tau\right] \leq \gamma$$

Define the critical region of $\tilde{T} : \mathbb{R}^{2d+1} \to \{0,1\}$ as:

$$K_{\tilde{T}} = \{(y, \theta, \sigma) \in \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty) : \tilde{T}(y, \theta, \sigma) = 1\}$$

Theorem 3. If $\tilde{T} : \mathbb{R}^{2d+1} \to \{0,1\}$ *is such that* $\tilde{T}(\cdot, \theta_0, \sigma_0)$ *has asymptotic level* γ *and constant conditional power function given* $\|\Theta - \theta_0\|_2 = \rho$ *for* $\mathbb{P}\|\Theta - \theta_0\|_2^{-1}$*-almost every* ρ *and if the critical region* $K_{\tilde{T}}$ is a $\mathbb{P}Z^{-1}$*-continuity set, then*:

$$\limsup_n \left(\Pi_n\left(\rho, K_{T_{\lambda_\gamma(\tau)}}\right) - \Pi_n(\rho, K_{\tilde{T}})\right) \geq 0 \qquad (2)$$

4. Experimental Results for Signal Detection in Non-asymptotic Regimes

Consider the detection problem (3), where $\Delta \in \mathrm{M}(\Omega, \mathbb{R}^d)$ is a bounded random interference and assume that we have a consistent estimator of σ. The standard Neyman-Pearson test and the GLRT cannot be used since the distribution of Δ is unknown. We can cast (3) in (1) with $\Theta = \varepsilon\theta_0 + \Delta$ and thus use $\tilde{T}_{\lambda_\gamma(\tau)}$ to perform the decision. The false alarm rate (FAR) of $\tilde{T}_{\lambda_\gamma(\tau)}$ tends to γ by upper values as N increases.

$$\begin{cases} \textbf{Observation: } Y = \varepsilon\theta_0 + \Delta + X, \text{ where:} \\ \quad \begin{cases} \varepsilon \in \{0,1\} \text{ is unknown} \\ \theta_0 \in \mathbb{R}^d \text{ is known, } X \sim N(0, \sigma^2 I_d) \\ \Delta \in M(\Omega, \mathbb{R}^d) \text{ has unknown distribution} \\ \|\Delta\|_2 \leq \tau \text{ for a known } \tau \geq 0 \\ \Delta \text{ and } X \text{ are independent} \end{cases} \\ \textbf{Testing problem: } \text{given } y = Y(\omega), \text{ test:} \\ \quad H_0 : \varepsilon = 0 \text{ vs. } H_1 : \varepsilon = 1 \end{cases} \qquad (3)$$

It is then possible to determine τ^* so that $\tilde{T}_{\lambda_\gamma(\tau^*)}$ maintains the FAR below γ for the detection problem (3). This adjustment is achieved as follows. Consider a uniform distribution for Δ on S_τ because this distribution is the least favorable in that it maximizes the FAR. Seek the value $k = 1, 2, \ldots$ such that the FAR of $\tilde{T}_{\lambda_\gamma(k\tau)}$ drops below γ. For the value of k thus found, fix $\tau_l = (k-1)\tau$ and $\tau_u = k\tau$. Then, by dichotomy, calculate $\tau^* \in (\tau_l, \tau_u)$ such that the FAR of $\tilde{T}_{\lambda_\gamma(\tau^*)}$ approximates at best γ without exceeding this level. In our experiments below, the dichotomy was stopped after 10 steps. The Monte-Carlo simulations were carried out with the following parameters: $d = 2$; $\sigma = 1$; $\tau = 1.77$, which corresponds to a Distortion-to-Noise ratio of 2 dB; $\|\theta_0\|_2 = 5.62$, which corresponds to an SNR of 12 dB; $\hat{\sigma}$ was the Maximum Likelihood Estimate from N iid standard Gaussian samples, with $N = \{20, 50, 100\}$. The ROC curves of $\tilde{T}_{\lambda_\gamma(\tau)}$ in Fig. 1 exhibit two important features. First, they are identical, which can be proved mathematically. Second, although $\tilde{T}_{\lambda_\gamma(\tau^*)}$ is not optimal in the sense of Theorem 3, it maintains the FAR below γ without much performance loss in comparison with the optimal $T_{\lambda_\gamma(\tau)}$ that requires a known σ.

5. Conclusion and Perspectives

In this paper, with respect to current issues in cybersecurity and ICS, we have presented Asymptotic RDT, which extends the initial RDT approach so as to take estimation of the model and the noise variance into account. We have illustrated the relevance of the approach through simulations and presented a way to compensate the effect of the estimation when detection is performed via Asymptotic RDT. The approach is very promising due to its genericity. Applications to change detection are in progress. Future works involve

extension to the case of an unknown noise covariance matrix and to noise distributions other than Gaussian.

References

[1]. V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, *ACM Computing Surveys*, Vol. 41, Issue 3, July 1 2009, pp. 1-58.

[2]. M. A. F. Pimentel, et al., A review of novelty detection, *Signal Processing*, Vol. 99, June 2014, pp. 215-249.

[3]. D. Pastor, Q.-T. Nguyen, Random distortion testing and optimality of thresholding tests, *IEEE Transactions on Signal Processing*, Vol. 61, Issue 16, Aug. 2013, pp. 4161-4171.

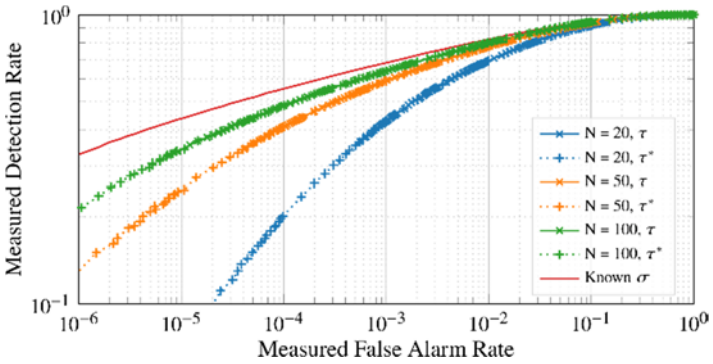


Fig. 1. ROC curves of $\tilde{T}_{\lambda_\gamma(\tau^*)}$ and $\tilde{T}_{\lambda_\gamma(\tau)}$.

(017)

Comparison Study of Two Recent Metaheuristic with Application to High Efficiency Induction Motor Design

H. Ladaycia
University of 20 Août 1955, Department of Electrical Engineering, Skikda, Algeria
Tel.: + 213795770990/+ 213676769423
E-mail: hania_ladaycia@live.fr

Summary: The design of the electrical machines is a complicated process. On top of that, it follows the objective physical and mathematical principles, but on the other hand, knowledge of these principles is often insufficient to produce a correct and economic design, and must be complete by considerable judgment, rules of thumb and experience which can be acquired only after intensive training on the job. This paper carries out efficiency optimization of 37 kW, 380 V, 1800 synchronous r.p.m., three phase squirrel cage induction motor (SCIM) using two swarm intelligence techniques Firefly algorithm (FA) and accelerated particle swarm optimisation (APSO). A comparison between APSO and FA swarm techniques is done. Efficiency and power factor are achieved to the goal and efficiency results satisfies high efficiency level (IE 2 class > 93.0 %). APSO is the best method for optimization in comparison with FA if efficiency is to be optimized.

Keywords: Metaheuristic, Design, Induction motor, High efficiency.

1. Introduction

Electric motors consume over half of the energy produced by power stations. This represents almost three-quarters of industrial consumption and almost half of the consumption of commercial sectors in industrialised countries. Motors therefore represent the most significant electrical loads and are therefore the main targets for achieving significant energy savings [1, 2].

The design of asynchronous motors covers a very wide field of activity, ranging from machines with a few watts of power intended for computer applications to a motor with a few megawatts installed in heavy industrial processes. This type of machine represents a large part of the electrical machinery market, especially squirrel-cage machines, are the most commonly used electrical machines [3]. Considerable efforts have been and still are being made to perfect the theory and methods of analysis and designs thereof, in order to improve its performance and optimise its price [1-15].

2. Proposed Method

In this work we present an optimal design method to optimise the three-phase induction motor in the manufacturing process. The APSO and FA have been adapted to optimise its design to improve its performances. The motor design procedure consists of a system of nonlinear equations imposing the characteristics of the induction motor, motor performances, magnetic stresses and thermal limits.

The results of the computer simulation are given to show the efficiency of the proposed design process and the optimisation techniques applied.

3. Problem Formulation

APSO and FA are implemented with the direct design described in [13] formulated by the following equation:

$$\begin{cases} \min \sum losses\left(D_{is}, d_1, d_2, h_r, b_{tr}, b_{ts}, b_{s1}, b_{s2}, h_s\right) \\ s.t. \quad \cos\phi \geq 0.861 \\ \quad \eta \geq 0.936 \\ \quad T_{c0} \leq 101 \\ \quad i_{LR} \leq 6 \\ \quad t_{LR} \geq 1.75 \\ \quad t_{bk} \leq 2.5 \\ \quad D_{out} = 343 \\ \quad D_{shaft} = 70 \end{cases}, \qquad (1)$$

where losses is the total losses in induction motor, D_{is} is the stator bore diameter, d_1 is the rotor slot higher diameter, d_2 is the rotor slot lower diameter, h_r is the rotor slot useful height, b_{tr} is the rotor tooth width, b_{ts} is the stator tooth width, b_{s1} is the stator slot lower width, b_{s2} is the stator slot higher width, h_s is the stator slot useful height, η is the efficiency, T_{c0} is the winding temperature, i_{LR} is the per unit locked rotor current, t_{bk} is the per unit breakdown torque, D_{out} is the stator outer diameter, D_{shaft} is the shaft diameter.

4. Simulation Results

The change of objective function values according to the iteration is shown in Fig. 1 it is seen that the APSO method converges at the smallest iteration number. The APSO algorithm converges at the 29th iteration. However, FA algorithm converges at the 127th iteration.

The performances and the optimal design variables of the motor obtained with the two methods are given in Table 1.

It is clear from the table that the APSO is able to offer the better efficiency and the desired power factor than that of the FA approach.

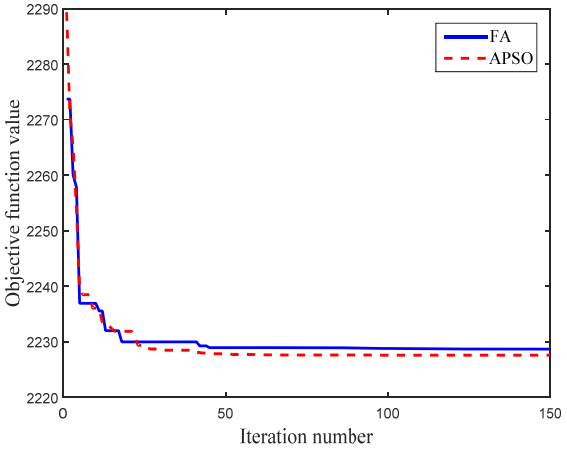


Fig. 1. Iteration number versus the objective function value.

Table 1. Comparison of the results obtained by the FA and APSO methods for a 37 kW SCIM.

Designation	Conception conventionelle	FA	APSO
D_{is} [mm]	221	215	215
d_1 [mm]	7.3477	7.5	7.5
d_2 [mm]	2.8419	2.5	2.5
b_{s1}	6.5195	5.2430	5.289712
b_{s2}	10.3856	8.7699	8.943088
h_r [mm]	28.6257	28.7	28.69998
b_{tr} [mm]	9.2754	5.0420	5
b_{ts}[mm]	8.2721	4.4	4.4
h_s [mm]	29.4923	28.3769	28.03786
i_{LR}	5.9638	5.9812	5.9959
t_{LR}	0.9928	0.9964	1.0008
t_{bk}	2.9135	2.9278	2.9323
Power factor [%]	86.44	86.18	86.18
Efficiency [%]	93.23	93.51	93.52
T° of winding [°C]	112.2971	113.2397	113.8508

5. Conclusions

APSO and FA are two powerful populations based heuristic algorithms for solving multimodal optimisation problems. The results of SCIM performances clearly demonstrate the ability of the APSO and FA to produce the best global design parameters that maximise its efficiency and starting torque as well. It has been exhibited that the proposed approaches encourage the continued use of APSO and will go a long way in serving as a useful tool in design problems.

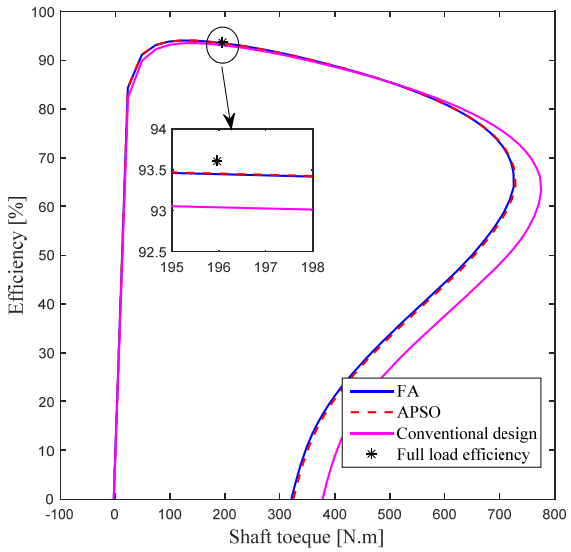


Fig. 2. Efficiency-Shaft torque characteristic after optimization.

References

[1]. R. N. Hasanah, A contribution to energy saving in induction motors, PhD Thesis, *EPFL*, 2005.

[2]. A. S. Sindekar, A. R. Agrawal, V. N. Pande, Comparison of some optimization techniques for efficiency optimization of induction motor, *International Journal of Engineering Science and Technology*, Vol. 5, Issue 6, 2013, 1303.

[3]. M. Ćalasan, M. Micev, Z. M. Ali, A. F. Zobaa, S. H. Abdel Aleem, Parameter estimation of induction machine single-cage and double-cage models using a hybrid simulated annealing-evaporation rate water cycle algorithm, *Mathematics*, Vol. 8, Issue 6, 2020, 1024.

[4]. R. Chaudhary, R. Sanghavi, S. Mahagaokar, Optimization of induction motor using genetic algorithm and GUI of optimal induction motor design in MATLAB, in Advances in Systems, Control and Automation, *Springer*, Singapore, 2018, pp. 127-132.

[5]. M. Çunkaş, R. Akkaya, Design optimization of induction motor by genetic algorithm and comparison with existing motor, *Mathematical and Computational Applications*, Vol. 11, Issue 3, 2006, pp. 193-203.

[6]. S. Mallik, K. Mallik, A. Barman, D. Maiti, S. K. Biswas, N. K. Deb, S. Basu, Efficiency and cost optimized design of an induction motor using genetic algorithm, *IEEE Transactions on Industrial Electronics*, Vol. 64, Issue 12, 2017, pp. 9854-9863.

[7]. A. S. Sindekar, A. R. Agrawal, V. N. Pande, Comparison of some optimization techniques for efficiency optimization of induction motor, *International Journal of Engineering Science and Technology*, Vol. 5, Issue 6, 2013, 1303.

[8]. L. Tutelea, I. Boldea, Induction motor electromagnetic design optimization: Hooke Jeeves method versus genetic algorithms,. in *Proceedings of the 12th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM'10)*, 2010, pp. 485-492.

[9]. C. Li, A. Rahman, Three-phase induction motor design optimization using the modified Hooke-Jeeves method,

Electric Machines and Power Systems, Vol. 18, Issue 1, 1990, pp. 1-12.

[10]. S. Subramanian, R. Bhuvaneswari, Comparison of modern optimization techniques with applications to single-phase induction motor design, *Electric Power Components and Systems*, Vol. 34, Issue 5, 2006, pp. 497-507.

[11]. D. Fodorean, L. Idoumghar, N'diaye, A., D. Bouquain, A. Miraoui, Simulated annealing algorithm for the optimisation of an electrical machine, *IET Electric Power Applications*, Vol. 6, Issue 9, 2012, pp. 735-742.

[12]. S. F. Contreras, C. A. Cortés, M. A. Guzmán, Bio-inspired multi-objective optimization design of a highly efficient squirrel cage induction motor, in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2016, pp. 549-556.

[13]. H. Ladaycia, A. Boukadoum, M. Mordjaoui, Firefly algorithm optimization of high efficiency induction machine using inverse problem, *International Review of Automatic Control*, Vol. 11, Issue 2, 2018, pp. 1974-6059

[14]. P. S. Prakash, P. Aravindhababu, Firefly optimization based design for improving efficiency of induction motor, *ARPN Journal of Engineering and Applied Sciences*, Vol. 10, Issue 4, 2015, pp. 1711-1716.

[15]. R. Chaudhary, R. Sanghavi, S. Mahagaokar, Optimization of induction motor using genetic algorithm and GUI of optimal induction motor design in MATLAB, in Advances in Systems, Control and Automation, *Springer*, Singapore, 2018, pp. 127-132.

(020)

The Pulse Project: A Framework for Supervising Data Exchanges in an IoT System

Jannik Laval [1]
[1] University Lumiere Lyon 2, DISP Laboratory, Lyon, France
Tel.: +33 478774306
E-mail: Jannik.laval@univ-lyon2.fr

Summary: This article describes our work in the supervision of data exchange in an IoT information systems. Traditional monitoring approaches and tools do not allow to apprehend the volume, speed and variety of alerts generated. Among these locks impacting service continuity, our work addresses the internal supervision of a distributed information system. In this work, we address the collection, modelling and processing of data in order to analyse and identify the evolution of the system architecture and the problems that may arise. The proposed framework supports various data exchange protocols, including AMQP and MQTT. It allows to navigate in the proposed structure, to query and visualize the architecture. Thanks to this framework, architecture problems can be identified. We have implemented the approach and tested it on different systems, including RabbitMQ. The next step will be to implement it for ROS.

Keywords: Data monitoring, IoT system, Dynamic analysis of IoT architecture, Interoperability assessment, Message broker.

1. Introduction

Event-driven architectures (EDAs), like the ones used in IoT, allow a loose coupling, but limit the ability to understand the overall behaviour of the architecture. The multiplicity of data exchanges generates complexity and highlights supervision needs that can be addressed by establishing monitoring and analysis systems [1]. A monitoring system is defined as a distributed process or set of processes that includes the function and dynamic collection, interpretation and processing of information about an application that is running.

Existing structures, including ROS [5], provide monitoring consoles. However, they mainly focus on low-level control information such as message frequency, performance indicators or memory usage.

The supervision of IoT systems involves integrating the dynamic aspect of the information systems architecture and the management of unpredictable phenomena, such as constituent security issues. The collection of information coming from the system on the one hand and from the technical vulnerabilities of the constituents on the other hand is a first step in keeping an IoT system in operation.

In this context, the main research problem addressed is to ensure the supervision of an IoT system based on the metadata of the messages transiting through the system.

2. The Pulse Framework

Our approach uses monitoring to perform an analysis capable of defining a kind of classification of the potential causes of problems occurring in exchanges in order of importance for a given problem. In a first study, we recommend that the behaviour of data exchange between communicating applications gives an indication of the level of interoperability of their data. A monitoring system should provide elements to maintain a good level of interoperability based on interoperability requirements [2].

We therefore propose to combine the information in order to perform advanced monitoring and interrogation. This would provide indicators for determining maintenance actions. The Pulse framework is composed of four levels: (i) A data import level, (ii) A time management and model version management level, (iii) A persistence level, (iv) An analysis level.

First, it collects the metadata of the messages transiting the IoT system in order to analyse its architecture. For this we proposed the analysis framework illustrated in Fig. 1. It collects metadata in order to (i) track the messages exchanged, (ii) simplify the visualization of the exchanges, (iii) improve maintainability by detecting exceptions.

Second, it organises the collected metadata so that they can be processed in a coherent manner. For this, we proposed a generic and extensible metamodel adapted to the EDAs. It is capable of analysing the AMQP, MQTT, KAFKA and CoAP protocols [3, 4]. We are currently working on adapting our metamodel to the ROS middleware (Robot Operating System is an open source system dedicated to robotics, now recognised as a standard software platform and widely used in the scientific community [5]). We have already theoretically valued the adequacy between ROS and the Pulse metamodel. It aims to represent three aspects of the information system: (i) a static representation: the architecture of the implemented system, representing the queues and channels that messages take; (ii) a dynamic representation, where messages are represented from the source component to the destination. It represents connections, messages, and

any volatile entities; (iii) the life-cycle representation of the architecture, where components are created, modified, deleted.

Third, it identifies interoperability problems and help resolve them. To do this, we have carried out a study of the causes and effects of the problems while identifying the indicators related to interoperability [3].

Fourth, it takes into account the life cycle of the different components. For this purpose, the Pulse structure integrates dynamic modelling functionalities, where the life cycle of the different components of the architecture is described including a history and the dates of creation and deletion of the components.

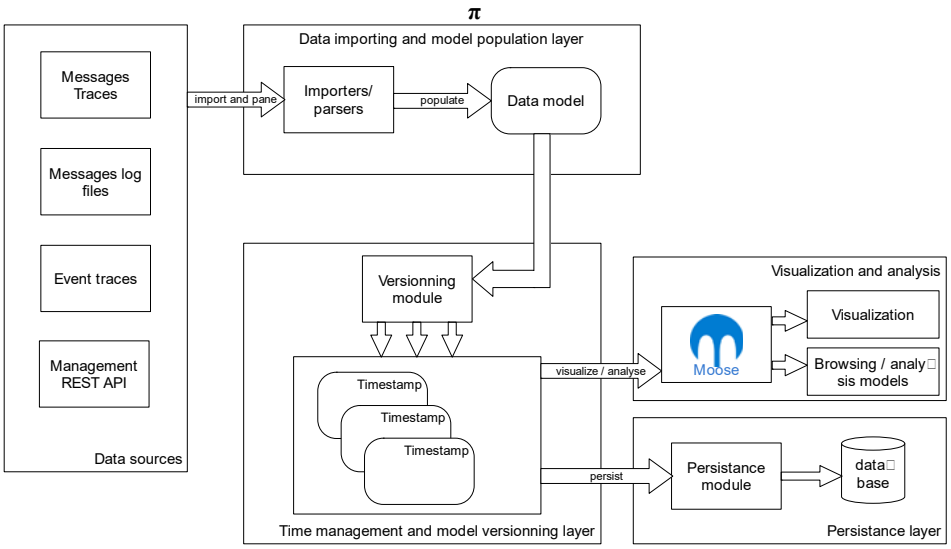


Fig. 1. The Pulse Framework [3].

3. Implementation

We set up a prototype tool of the Pulse framework. It allows us to visualize some of the system's behaviours. The objective of this prototype is to allow to understand the status of each message and to provide indications facilitating the analysis of interoperability problems. The prototype is implemented on top of Moose, an open source data analysis platform [6, 7], in which we have implemented the Pulse metamodel. The extension of Moose is a major asset for our implementation approach, as it allows us to build on existing tools and reuse them.

The prototype shows the feasibility and interest of our approach. Its implementation is composed of all parts of the framework (Fig. 1): the Pulse metamodel, the importer development, the versioning management thanks to Orion [8], the persistence layer, and some visualisations.

4. Conclusion

In this article, we presented Pulse Framework. This framework is already used in AMQP type systems [3, 4]. The metamodel used by our framework is generalizable and we are currently working on integrating the messages transiting on the ROS middleware.

The next step is to automatically collect the data transiting between the ROS nodes and provide analysis tools (visualisation, metrics) to identify problems that may appear at runtime.

References

[1]. T. Brand, H. Giese, Generic adaptive monitoring based on executed architecture runtime model queries and events, in *Proceedings of the IEEE 13th International Conference on Self-Adaptive and Self-Organizing Systems (SASO'19)*, June 2019, pp. 17-22.

[2]. S. Mallek, N. Daclin, V. Chapurlat, Towards a conceptualization of interoperability requirements, in Enterprise Interoperability IV, *Springer*, 2010, pp. 439-448.

[3]. J. Laval, N. Amokrane, N. Moalla, M. Derras, Analysis of data exchange among heterogeneous IoT systems, in *Proceedings of the 10th International Conference on Interoperability for Enterprise Systems And Applications (IESA'20)*, Tarbes, France, March 24-27, 2020.

[4]. N. Amokrane, J. Laval, P. Lanco, M. Derras, N. Moala, Analysis of data exchanges, contribution to data interoperability assessment, in *Proceedings of the International Conference on Intelligent Systems (IS'18)*, Sep. 2018, pp. 199-208.

[5]. M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y Ng, ROS: An open-source Robot Operating System, in *Proceedings of the ICRA Workshop on Open Source Software*, Vol. 3, Kobe, Japan, 2009.

[6]. S. Demeyer, S. Tichelaar, S. Ducasse, FAMIX 2.1 – The FAMOOS Information Exchange Model, Rapport Technique, *University of Bern*, 2001.

[7]. S. Ducasse, T. Gîrba, A. Kuhn, L. Renggli, Meta-environment and executable meta-language using Smalltalk: An experience report, *Journal of Software and Systems Modeling (SOSYM)*, Vol. 8, Issue 1, Feb. 2009, pp. 5-19.

[8]. J. Laval, S. Denier, S. Ducasse, J.-R. Falleri, Supporting simultaneous versions for software evolution assessment, *Science of Computer Programming*, Vol. 76, Issue 12, 2011, pp. 1177-1193.

(022)

EEG Based BCI System for Driver's Arm Movements Identification

E. Zero, C. Bersani and R. Sacile
Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova,
Via all'Opera Pia 13, 16145 Genova, Italy
E-mail: enrico.zero@dibris.unige.it

Summary: Automation systems in the automotive context are moving towards new challenging applications. Although the state of the art of these technologies is advanced, the presence of the human component in such systems still remains a critical issue in terms of safety. Autonomous vehicles represent an example where the transition to a fully automated vehicle still requires the presence of the human which can represent the failure or added value of such systems. The possibility to monitor driver's conditions becomes crucial to detect common or unexpected driver behaviour and to identify, if any, accidents in advance. In this paper, an electroencephalogram (EEG)-based Brain-Computer Interface (BCI) system is proposed to detect and identify the driver's brain activities during common driving task. The aim is to explore brain electrical signals when a driver has to rotate the steering wheel to perform a right or a left turn in order to execute a curve on a path. A time delay neural network (TDNN) technique has been implemented to classify the human's EEG signals when the participant has to drive a car along a curvilinear path in a virtual simulated scenario. Results demonstrated that the TDNN models may represent a suitable tool to identify driver's arm movements but further efforts have to be dedicated to improve the quality of recognition.

Keywords: Electroencephalogram (EEG), Brain-Computer Interface (BCI) system, Movements detection and recognition.

1. Introduction

In the road transport chain, the safety aspect represents one of the main concern. Automated driving, assistant driver systems, and platooning have been recently introduced in automotive sector [1]. Although various types of automation are adopted, these technologies are not the unique solutions addressing negative road safety results, in fact, the driver still remains one of the most critical elements in terms of accident causes. In this context, the main challenge is to provide control vehicle system with the possibility to identify autonomously critical situation during a tour and driver's intentions while performing driving tasks [2]. Considering the driver as an integral component of the intelligent vehicle automated system, the objective is to monitor his/her behaviour in order to acquire data and information to enhance the automated system's performance, to adjust unforeseen errors, and support the decision making process.

The emerging technology which is able to create a direct communication channel between the human brain and external devices is the Brain-Computer Interface (BCI). The BCI offers the potential to predict the intention of the monitored human by analysing the electrical brain signal and to transit control input to the external devices. The electroencephalography (EEG) represents, at the present, the most reliable brain signal acquisition system. EEG provides the electrical brain activities by monitoring the voltage variation through electrodes allocated on the scalp. Many emotional and cognitive processes may be detected by physiological signals EEG based methods [3].

In the driving context, the EEG signal recognition has been applied to identify specific driver's mental state such that driving fatigue [4], slipness [5], aletness or distraction [4]. In general, in those approaches, the partecipant is subjected to external stimuli called evoked potentials in order to perceive changes in heart rate [7] or eyes blinking [8]. In [9], the authors demonstrated that unexpected acoustic alarms and visual external events may be strongly correlated to the drivers' brain activities during a simulated driving sessions.

More complex is the task to identify the imaginary or the real execution of the body movements by EEG signals variations [10]. Large literature is dedicated to identify the main reliable classifier to correlate the movement related EEG signal into the corresponding actual body movement. Different kind of classifier are presented in this context: learning vector quantization (LVQ) for hand movement [11], support vector machine (SVM) for arms [12] or fingers movements [13]. The authors in [14] proposed widespread reviews on EEG signal classification.

In this paper, the behaviour of a driver who turns the steering wheel of the vehicle while preforming a curvilinear path in a virtual driving environment is analysed. The objective of this study is to detect and identify, by the driver's brain activities, the different movements of the arms to turn and take a left or right curve in a path. A time delay neural network (TDNN) classification model was built to predict right and left turns during actual driving. Different experiments have been carried out involving three participants, with driving license.

2. Material and Methods

2.1. Enobio EEG cap

An Enobio Cap 8 Features was adopted in the experiments. Signals were recorded from six different channels, namely F7, FZ, F8, C4, C3, CZ, in accordance with the international 10-20 system. The selected channels are shown in the Fig. 1.

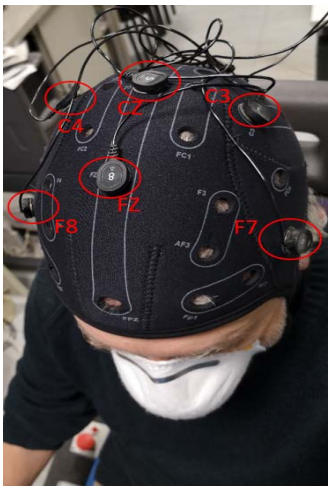


Fig. 1. Electrodes placement in the ENOBIO cap. (Red circles represent the electrodes monitored during the experiments).

The electrodes are located in the frontal and central lobes as this part of the brain controls the complex behavior including the body awareness, spatial orientation and attention.

2.2. Test Drive

Each driver test has been performed three times by a single participant who was 30 years old with driving license. The participant was seated on a common chair in front of a driving simulator. This latter consisted in a steering wheel connected to the pedal system, and to a LCD screen on which the simulation environment was projected. The driver was able to rotate the steering wheel, to accelerate or brake as if he was actually driving a real car. The simulation lasted 15 minutes and, the user had to drive along a tour represented by the figure eight route as shown in the Fig. 2. OKTAL Scanner Studio is the software used to generate the simulation environment. The focus of this paper is to study the correlation between the arm movements of the driver while performing the right/left curves on the road and the related recorded EEG signals.

2.3. Elaboration Data

The elaboration data was performed in Matlab R2020b. A high pass filter to 0.167 Hz was applied to the EEG signals to remove the direct current shift. Three different analysis have been realized by the implementation of time delay neural network (TDNN). TDNN has been used due to its capability to generate a finite dynamic response to time series input data.

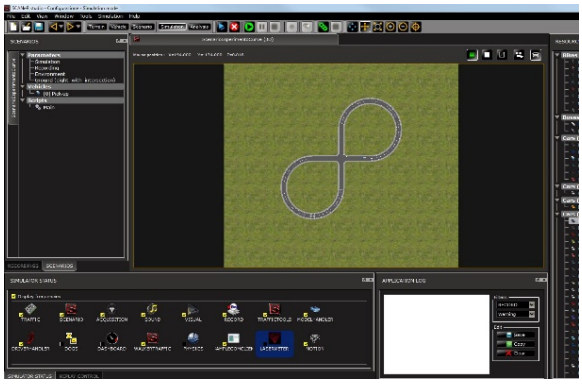


Fig. 2. Test drive circuit in OKTAL Simulator environment.

The TDNN is typically successfully used in phoneme recognition but its large applications in EEG signal analysis [15] has been verified due to the non – stationary nature of the brain activities. In [16] the TTDN is applied for the classification of finger movements with a recognition rate of 93.02 %.

The TDNN performed in this paper, whose structure is presented in the Fig. 3, has 10 time delays and 4 hidden fields.

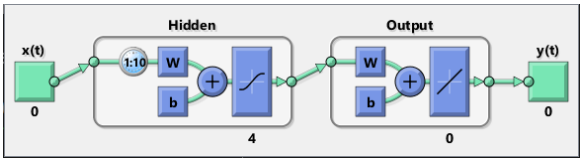


Fig. 3. TDNN used in the paper.

2.4. Experiments

The experiment involved 3 measurement sessions, once in three different days. The data related to the first session is indicated by number 1, the second as number 2 and the last one as number 3.

Three different analysis has been carried out on the monitored EEG signals.

The first test is related to the classifier generation by the Levenberg Marquardt algorithm based TDNN. The data acquired in each session was divided in two equal number of samples, labelled as 1A, 1B, 2A, 2B, 3A, 3B. The different sets of data have been used to train and test six TDNNs in order to identify the best classifier in term of correlation. In Table 1, the different training and test sets used in the applications appear. The main objective of the first analysis is to evaluate if the increasing number of input data may reflect a better recognition performance.

In the second trial, the TDNNs, realized in the previous test, have been modified. In particular, the NN4, NN5 and NN6 has been revised by modifying the cost function which minimizes the Mean Square Error (MSE). In this second approach, a weight p_i has been introduced, for each sample which belongs to the training set, in the objective function as in (1).

$$MSE = \frac{1}{N*S}\sum_{k=1}^{S}\sum_{i=1}^{N} p_i^s \, (y_i^s - \hat{y}_i^s)^2, \qquad (1)$$

where MSE is the mean square error; S is the number of dataset s-th used in the training; N is the number of elements in each dataset s-th used in the training set; p_i^s is the weight associated to each data i-th belonging to the dataset *s-th*; y_i^s is the observed value of the output in the dataset *s-th*; $\hat{y}_i^s$ is the predicted value for the output in the dataset *s-th*.

Table 1. Set of data used to train and test the TDNNs.

TDNN	Training set	Test set
NN1	1A	1B
NN2	2A	2B
NN3	3A	3B
NN4	2A, 2B, 3A	3B
NN5	1A, 1B, 2A, 2B, 3A	3B
NN6	1A, 1B, 2A	2B

The new TDNNs, namely NN7, NN8 and NN9, are trained and tested by the same data set already adopted, respectively, in NN4, NN5 and NN6. However, in this new approach, the objective functions for the training phase have been modified balancing the different input components as appear in Table 2. Note that the data coming from the older tests drive have a lower weight in respect to the recent ones.

Table 2. TDNNs characteristics.

TDNN	Dataset *s-th* in the objective function	Weights p_i^s	Test set
NN7	2A	0.5	3B
	2B	0.5	
	3A	1	
NN8	1A	0.25	3B
	1B	0.25	
	2A	0.5	
	2B	0.5	
	3A	1	
NN9	1A	0.5	2B
	1B	0.5	
	2A	1	

The aim of the second analysis is related to identify the set of parameters which, by multiplying the objective components, increase the accuracy of the recognition.

The third test is carried out to evaluate TDNN performance variations according to different values of the weights, calibrated or random, associated to input-output relationship function of the network. In particular, the performances have been evaluated comparing the recognition accuracy and the computation time, for the TDNN initialized randomly by the system and the TDNN where we actually updated each weight on corresponding connections with the same values obtained as the output of the training phase according to the set of data which appears in the Table 3.

Table 3. Generation of the weights used in the TDNNs.

TDNN	Weights initialization	Training set	Test set
NN10	weights generated by 1A, 1B, 2A, 2B	3A	3B
NN11	weights generated by 1A,1B	2A	2B
NN12	random weights	1A, 1B	2A
NN13	weights generated by 1A	1A, 1B	2A
NN14	random weights	1A, 1B, 2A	2B
NN15	weights generated by 1A and 1B	1A, 1B, 2A	2B
NN16	random weights	1A, 1B, 2A, 2B	3A
NN17	weights generated by 1A,1B, 2A	1A, 1B, 2A, 2B	3A
NN18	random weights	1A, 1B, 2A, 2B, 3A	3B
NN19	weights generated by 1A,1B, 2A, 2B	1A, 1B, 2A, 2B, 3A	3B

3. Results

Table 4 shows the R index of each TDNN in the first analysis. The results show a significant correlation between the EEG signal and the actual arm movements of the participant. The average value is R = 0.74 and, in three cases, R > 0.8.

Table 4. First analysis results.

Simulation	R	MSE
NN1	0.8159	0.3657
NN2	0.8538	0.2828
NN3	0.6079	0.5130
NN4	0.6486	0.4528
NN5	0.6588	0.4622
NN6	0.8507	0.292

However, by these results, a significant improvements of the performances related to the increasing number of input data is not evident. Fig. 4 display the predictive and observed values of NN1.

In the second analysis, a strong correlation is verified, too. However, the modified objective function generates better values, both for R and MSE, only in the NN9 in respect to the NN6.

In the third analysis, the value R is greater than 0.5 in the 80 % of cases, and the MSE is lower than 0.5 in the 50 %. On the contrary, in respect to the recognition accuracy for TDNNs initialized randomly or with accurate weights input values, the results appear very similar in the two approaches. Besides, from the computation viewpoint, the TDNNs with weight initialization appears, in general, less time expensive in respect to the ones with random initialization.

Table 5. Second analysis results.

Simulation	R	MSE
NN7	0.6408	0.4633
NN8	0.6442	0.4464
NN9	0.8502	0.2892

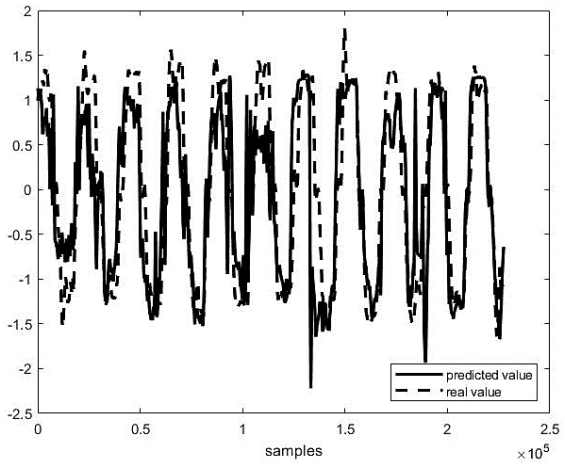


Fig. 4. Predictive and real value in NN1.

Table 6. Third analysis results.

Simulation	R	MSE	TIME (sec)
NN10	0.6040	0.4949	100.052
NN11	0.8517	0.2864	24.882
NN12	0.4501	0.8698	114.260
NN13	0.4696	0.7928	54.245
NN14	0.8542	0.2854	196.969
NN15	0.8495	0.2912	567.360
NN16	0.5545	0.6188	442.865
NN17	0.5983	0.5841	127.593
NN18	0.6056	0.5841	209.203
NN19	0.6392	0.4781	644.334

4. Conclusions

The main objective of this work is to tackle the problem of driver's arms movements detection and recognition using EEG signals and to classify the acquired data with a TDNN classification algorithm. Results demonstrated a statistically significant positive correlation of the EEG signals with the actual participant's motions to rotate the steering wheel to compute a curve in the virtual driving simulation environment. Other more detailed studies have to be implemented to verify the improvements in the recognition accuracy taking into account an increasing number of dataset for the training phase and for the weight parameters initialization in the objective function and in the input output function of the TDNNs.

References

[1]. S. Graffione, B. Bersani, R. Sacile, E. Zero, Model predictive control for cooperative insertion or exit of a vehicle in a platoon, in *Proceedings of the 17th International Conference on Informatics in Control, Automation and Robotics (ICINCO'20)*, 2020, pp. 352-359.

[2]. S. Graffione, C. Bersani, R. Sacile, E. Zero, Model predictive control of a vehicle platoon, in *Proceedings of the IEEE 15th International Conference of System of Systems Engineering (SoSE'20*), 2020, pp. 513-518.

[3]. N. K. Al-Qazzaz, M. K. Sabir, S. H. B. M. Ali, S. A. Ahmad, K. Grammer, Electroencephalogram profiles for emotion identification over the brain regions using spectral, entropy and temporal biomarkers, *Sensors*, Vol. 20, Issue 1, 2020, 59.

[4]. P. Shangguan, T. Qiu, T. Liu, S. Zou, Z. Liu, S. Zhang, Feature extraction of EEG signals based on functional data analysis and its application to recognition of driving fatigue state, *Physiological Measurement*, Vol. 41, 2020, 125004.

[5]. Y. Jiao, Y. Deng, Y. Luo, B. L. Lu, Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks, *Neurocomputing*, Vol. 408, 2020, pp. 100-111.

[6]. S. P. Kumar, J. Selvaraj, R. Krishnakumar, A. Sahayadhas, Detecting distraction in drivers using electroencephalogram (EEG) signals, in *Proceedings of the Fourth International Conference on Computing Methodologies and Communication (ICCMC'20)*, 2020, pp. 635-639.

[7]. S. De Nadai, M. D'Incà, F. Parodi, M. Benza, A. Trotta, E. Zero, R. Sacile, Enhancing safety of transport by road by on-line monitoring of driver emotions, in *Proceedings of the 11th System of Systems Engineering Conference (SoSE'16)*, 2016, pp. 1-4.

[8]. G. Li, W. Y. Chung, Estimation of eye closure degree using EEG sensors and its application in driver drowsiness detection, *Sensors*, Vol. 14, Issue 9, 2014, pp. 17491-17515.

[9]. E. Zero, C. Bersani, L. Zero, R. Sacile, Towards real-time monitoring of fear in driving sessions, *IFAC-PapersOnLine*, Vol. 52, Issue 19, 2019, pp. 299-304.

[10]. R. Chaisaen, P. Autthasan, N. Mingchinda, P. Leelaarporn, N. Kunaseth, S. Tammajarung, T. Wilaiprasitporn, Decoding EEG rhythms during action observation, motor imagery, and execution for standing and sitting, *IEEE Sensors Journal*, Vol. 20, Issue 22, 2020, pp. 13776-13786.

[11]. G. Pfurtscheller, J. Kalcher, C. Neuper, D. Flotzinger, M. Pregenzer, On-line EEG classification during externally-paced hand movements using a neural

network-based classifier, *Electroencephalography and clinical Neurophysiology*, Vol. 99, Issue 5, 1996, pp. 416-425.

[12]. D. Planelles, et al., Evaluating classifiers to detect arm movement intention from EEG signals, *Sensors*, Vol. 14, Issue 10, 2014, pp. 18172-18186.

[13]. X. Liao, D. Yao, D. Wu, C. Li, Combining spatial filters for the classification of single-trial EEG in a finger movement task, *IEEE Transactions on Biomedical Engineering*, Vol. 54, Issue 5, 2007, pp. 821-831.

[14]. A. R. Chakole, P. V. Barekar, R. V. Ambulkar, S. D. Kamble, Review of EEG signal classification, in *Proceedings of the International Conference on Information and Communication Technology for Intelligent Systems (ICTIS'18)*, 2018, pp. 105-114.

[15]. R. Rao, R. Derakhshani, A comparison of EEG preprocessing methods using time delay neural networks, in *Proceedings of the 2nd International IEEE EMBS Conference on Neural Engineering (NER'05)*, 2005, pp. 262-264.

[16]. G. Emayavaramban, S. Ramkumar, A. Amudha, K. S. Kumar, Classification of hand gestures using FFNN and TDNN networks, *International Journal of Pure and Applied Mathematic*s, Vol. 118, Issue 8, 2018, pp. 27-32.

(023)

Robotic Manufacturing Systems Using Internet of Things: New Era of Facing Pandemics

Hamed Fazlollahtabar [1]
[1] Department of Industrial Engineering, School of Engineering, Damghan University, Damghan, Iran
Tel.: + 989111137298, fax: + 981132190118
E-mail: h.fazl@du.ac.ir; hfazl@iust.ac.ir

Summary: Robotics made a tremendous transformation in manufacturing by facilitating manufacturing operations. By evolution of Industry 4.0 and its related technologies, a new trend of digitalization in advanced manufacturing systems occurred. Internet of Things (IoT) in manufacturing environments, enables manufacturing related devices to communicate with each other and with the control unit to send data and receive manufacturing rules accordingly. Cooperation between robots and IoT provides new opportunities to streamline the production processes in an efficient and cost-effective manner. The integration of robotic agents with IoT provide the novel concept of the Internet of Robotic Things developing new possibilities in various industrial fields specifically in facing new pandemics. Human operators are not efficient anymore in pandemics leading to stoppage of production systems and a huge amount of lost sale, back orders, and economic loss. Here we emphasize on the new era of IoT robotic manufacturing in pandemics and inauguration of modern aspects modeling and decision making.

Keywords: Internet of things (IoT), Robotic manufacturing and assembly, Decision making.

1. Introduction

Internet of Things (IoT) has been extracted as a useful tool in Industry 4.0. IoT in manufacturing systems enable effective process managers to monitor and supervise the production using dispatching rules. One of the most efficient production systems is robotic one. Robots are able to process manufacturing tasks faster and with higher quality without working time constraint. The challenge is in the robot control system leading to real time decisions. To handle large amounts of data in manufacturing floor and production robot real time control, IoT is employed to deliver a mechanism through internet oriented technologies. Robotic manufacturing system is highly flexible to satisfy customized production according to customers' demands. Automation is cost-effective due to higher throughput and productivity. Thus, in this paper we aim to propose a comprehensive model to control IoT robotic manufacturing in pandemics to prevent lost sale and economic loss and to keep the sustainable production [1].

The interrelationships of modern manufacturing systems with the promotion of information technology (IT) related innovations leading to prompt data-driven decision making in all levels of a hierarchical production system based on real-time data exchange systems (Dumitrache and Caramihai, 2010). In the literature, the development of manufacturing systems is widely correlated with the advancement of IT; for instance computer numerical control (CNC) cooperated with industrial robots evolved flexible manufacturing systems (FMSs) and etc. Adoption of IT software development services convinced the enterprises to promote their conventional manufacturing systems to a higher level with respect to the IT infrastructure provided [2, 3].

The concept of smart manufacturing system prompt within Industry 4.0 paradigm denoting on employment of cyber-physical systems and internet of things to coordinate the manufacturing functions and transfer required data for appropriate and effective decentralized decision making. The communication, cooperation and coordination among entities themselves and human operators are the major key performance criteria to evaluate the success factors of the system [4].

Robotics and IoT are communication within the smart manufacturing to configure the Internet of Robotic Things (IoRT) paradigm. In IoRT, intelligent robots can sense the occurrences around a certain domain of distance and transfer through internet to a control and decision processing unit and distribute the data among other entities and then decide about the appropriate action.

2. Decision Autonomy in Robotic Manufacturing

Decisional autonomy refers to the ability of the system to determine the best course of action to fulfil its tasks and missions. This is mostly not considered in IoT middleware: applications just call an actuation API of so-called smart objects that hide the internal complexity. Roboticists often rely on Artificial Intelligence (AI) planning techniques based on predictive models of the environment and of the possible actions. The quality of the plans critically depends on the quality of these models and of the

estimate of the initial state. In this respect, the improved situational awareness that can be provided by an IoT environment can lead to better plans. Human-aware task planners use knowledge of the intentions of the humans inferred through an IoT environment to generate plans that respect constraints on human interaction. IoT also widens the scope of decisional autonomy by making more actors and actions available, such as controllable elevators and doors. However, IoT devices may dynamically become available or unavailable, which challenges classical multi-agent planning approaches. A solution is to do planning in terms of abstract services, which are mapped to actual devices at runtime.

With the fourth industrial revolution, where robotics, CPS, and cloud technologies are merged, different domains are benefiting from rapid development. In this context, the IoRT systems can provide several advantages over traditional robotic applications, such as offloading computation-intensive tasks on the cloud, accessing large amounts of data, and sharing information with other robots, aiming to learn new skills and knowledge from each other. Moreover, IoRT applications can also be used remotely, facilitating the work of both researchers and industrial operators, and making it more accessible, allowing cooperation between humans even from long distances.

Robot-based production represents the backbone of smart manufacturing, and the concept of industrial robots has been occurring a continuous change in the latest years, mainly due to the embedding of IoT technologies. IoRT represents the main enabler of such change, through which manufacturing is embracing the concepts of Industry 4.0, embedding sensors, automation, and monitoring of products and processes. The Fourth Industrial Revolution has transformed how products are manufactured, adapting to such technological innovation, aiming to produce high-quality goods and services. Smart manufacturing involves system flexibility, monitoring, and adaptation to change. Specifically, additive manufacturing is a critical process in terms of manufacturing methods. In fact, innovations in digital technologies that are occurring during the Fourth Industrial Revolution need to keep up with advancements in manufacturing processes and materials. In this scenario, with the aim of facilitating smart manufacturing by sensor systems, flexible electronics of additive manufacturing and their reliability during processes are of critical consideration.

3. Comprehensive IoT-Robotic Manufacturing System in Pandemics

Application of integrated robotics and IoT lead to enhance the technologies and opportunities related to healthcare specifically in facing pandemics such as the novel Coronavirus. The major infection way of pandemics is contact between people while IoRT provides a contactless technology to handle production tasks (or even in other areas such as service, agriculture, medical treatments, etc.) leading to a safer environment. Meanwhile, the production tasks are not stopped anymore while the whole process is performed and monitored by intelligent robots [5].

The COVID-19 pandemic implies the vulnerability of workers to infection causing tremendous loss in production, services, customer relationship management and all other enterprise functional modules. IoRT could handle all manufacturing tasks and perform measurement and collect data and information sensed by the sensors to facilitate intelligent decision making.

IoRT includes a triple intelligent modules:

1: Sensors that are used by robots to sense their environment to monitor the manufacturing floor and transfer data with other intelligent devices.

2: Analytical core which is able to mine the data collected and process using computational capabilities like edge computing (local computation) rather than cloud preventing huge amount of data transfer.

3: Intelligent action, that is, robot can act based on the collected data and analysis leading to a complete cyber-physical process where cooperation between machin2machine and man2machine emerges new services like IoRT-maintenance, IoRT-quality control, IoRT-inventory control and etc.

Robots are appropriate substituents of human workforce with respect to increase in labor price. Artificial intelligence is a plus in robotic systems to keep a standard production mode and required quality during the production cycle time. The process of man2machine boosts productivity with fewer number of workers.

IoRT is caple to face pandemics keeping a certain standard of manufacturing requirements and on the other hand provide remote man2machine communication through internet for control and coordination purposes promising appropriate multi-dimension data analysis and decision making. (See Fig. 1).

4. Conclusions

The IoRT emerges novel paradigm in modern manufacturing systems. Cooperative learning based on the sensed data provide a collective knowledge being employed by robots to handle their allocated tasks in an intelligent manner. IoT sensors help in collecting data and transmitting through a safe and fast route for further controlling or decision making tasks. Motion path planning and other controlling behaviors and actions of robots are based on the data gathered by sensors. IoT focuses on pervasive sensors while robots uses interactive ones to handle automated production and execute production tasks in any situations. Robots can perform their tasks in pandemics where human work force are not able to cooperate. This capability provide a huge opportunity to producers to absorb a larger amount of market share in competition.

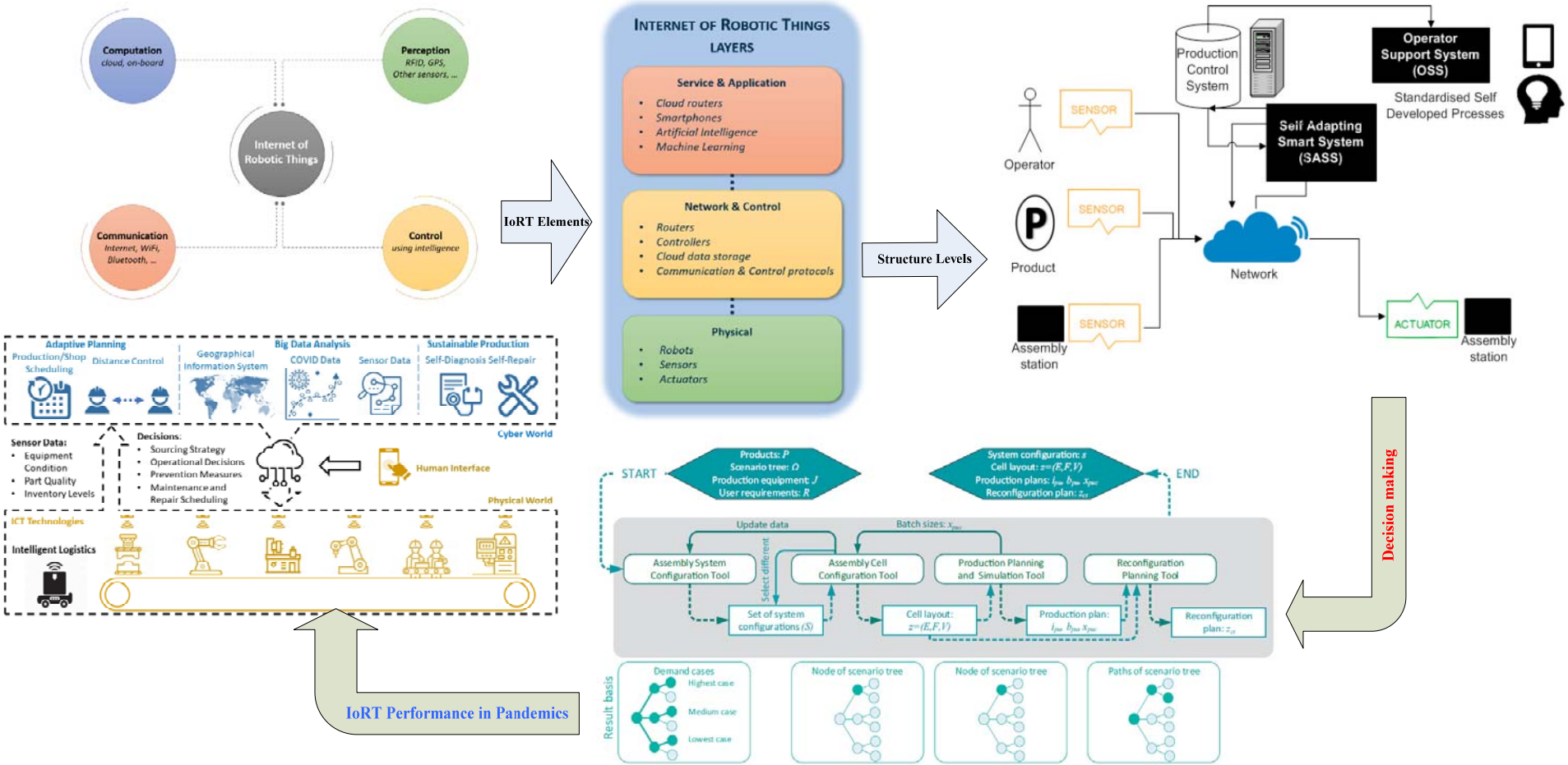


Fig. 1. Comprehensive IoT-Robotic Manufacturing System in Pandemic.

References

[1]. V. Alcácer, V. Cruz-Machado, Scanning the Industry 4.0: A literature review on technologies for manufacturing systems, *Eng. Sci. Technol. Int. J.*, Vol. 22, 2019, pp. 899-919.

[2]. C. Ju, H. I. Son, Modeling and control of heterogeneous agricultural field robots based on Ramadge-Wonham theory, *IEEE Robot. Autom. Lett*., Vol. 5, 2020, pp. 48-55.

[3]. M. Mahbub, A smart farming concept based on smart embedded electronics, internet of things and wireless sensor network, *Internet Things*, Vol. 9, 2020, 100161.

[4]. G. Valecce, G. Micoli, P. Boccadoro, A. Petitti, R. Colella, A. Milella, L. Alfredo Grieco, Robotic-aided IoT: Automated deployment of a 6TiSCH network using an UGV, *IET Wirel. Sens. Syst*., Vol. 9, 2019, pp. 438-446.

[5]. K. Raj Kumar, C. S. Kumar, C. Yuvashree, M. S. Murugan, Portable surveillance robot using IoT, *Int. Res. J. Eng. Technol. (IRJET)*, Vol. 6, 2019, pp. 94-97.

(024)

A Modified Reinitialization Mechanism for Particle Swarm Optimization Based Control, Case Study: PV System

T. Shaqarin [1]
[1] Tafila Technical University, Department of Mechanical Engineering, 66110 Tafila, Jordan
Tel.: + 962 3 2250326
E-mail: tshagareen@ttu.edu.jo

Summary: This paper proposes a modified reinitialization technique for online Particle Swarm Optimization (PSO). The need for reinitialization is motivated by the variation of the global maximum of the objective function due to varying operating conditions. A photovoltaic (PV) system is considered a convenient benchmark for maximum power point tracking (MPPT). The modified reinitialization technique involves monitoring the current global maximum to detect significant variation and the swarm positions to ensure convergence. Moreover, a partial reinitialization option is introduced as an intermediate conditional step between detection and full reinitialization. The simulation shows that partial reinitialization can improve convergence and efficiency.

Keywords: PSO, Reinitialization, MPPT, PV, Optimization-based control.

1. Introduction

Particle Swarm Optimization is a metaheuristic optimization technique that mimic the behavior of natural systems, such as: bird flocks, fish schools, and animal herds. The first milestone paper regarding PSO is written by Kennedy et al in 1995 [1]. Subsequently, there was a surge in the number of scientific papers attributed to the modification of PSO algorithm and swarm intelligence. The PSO algorithm is easily implemented, and it has a fast convergence, and it is used in solving multidimensional, multiobjective, and constrained problems [2].

Thus it has been applied in various applications that requires optimization, control, machine learning, etc.

1.1. PSO Algorithm

The algorithm of PSO randomly initializes a swarm of particles that represent a potential solution to an optimization problem in the search space. Then, positions of the particles are iterated with the aim of converging to a global optimum of an objective function. The algorithm then monitors the positions of each particle, and records the best solution for each particle (personal best p). Furthermore, PSO records the best solution over the swarm (global best p_g) at each iteration. For each future iteration, the position (x) and velocity (v) of the particles are computed as a function of the best position of the swarm (social component), the best personal position of the particle (cognitive component), and its previous velocity (memory component). A PSO has many variants, in this work, a PSO with damped inertial weight (PSO-DIW) is used as represented by the following group of equations:

$$\begin{aligned} v_{ij}(t+1) &= w v_{ij}(t) + c_1 R_1 (p_{ij}(t) - \\ &- x_{ij}(t)) + c_2 R_2 (p_{g,j}(t) - x_{ij}(t)), \end{aligned} \quad (1)$$

$$\begin{aligned} x_{ij}(t+1) &= x_{ij}(t) + v_{ij}(t+1) \\ i &= 1,2,\dots,N \; j = 1,2,\dots,n, \end{aligned} \quad (2)$$

$$w = w \times w_{damp}, \quad (3)$$

where N is the number of particles, n number of dimensions. Coefficient w is the inertial weight, and the product of w with velocity of a particle represents the momentum of the particle. Coefficients c_1 and c_2 are called the acceleration coefficients and R_1 and R_2 are randomly generated, model specific numbers between the range of 0 to 1.

1.2. Reinitialization Mechanism

Reinitialization is conceivably the less studied aspect of the PSO algorithm compared to its other modifications. However, reinitialization is required due to varying operating conditions leading to a change in the global maximum of the objective function.

A multitude of significant reinitialization methods are mentioned in the literature. Miyatake *et al* [3] have modified the algorithm by imposing two constraints that detects the solver convergence and the global best change. Here convergence is detected by the change in velocity, which could be misleading when imposing a large damping term in the PSO algorithm. Romasevich and Loveikin [4] introduced a global best rate as criterion for the PSO reinitialization, while they ignored PSO convergence. Martínez-Cruz et al [5] presented a mechanism that detect the global best change, but the reinitialization is activated after specific number of evaluations. This technique require a prior knowledge of the process that can identify the number of evaluations needed for convergence. Rout et al [6] introduced reinitialization based on the global best change using two past records, while ignoring

convergence of the PSO. Eltamaly *et al* [7] proposed a reinitialization strategy that search for new global best by sending a particle to anticipated locations of them. This method has the advantage of avoiding unnecessary reinitialization, but again, it needs a prior knowledge of the system.

In this work an alternative, online, initialization mechanism is proposed that circumvent the drawbacks of previous implementations such as ignoring the need of PSO convergence, and the requirement of a great level of prior knowledge for the process.

The proposed mechanism introduces two constrains as shown in the following equations:

$$\|x_{ij}\|_2 < A, \tag{4}$$

$$\frac{|p_{gj(t+1)} - p_{gj(t)}|}{p_{gj(t)}} > B, \tag{5}$$

where both A and B represent the constraints (to be selected). When conditions in equations 4 and 5 are satisfied, respectively, the swarm is reinitialized. Condition (4) ensures that all the particles converge to almost the same position, as close as possible, depending on the value of A. While B is controlling the global best rate to reinitialize. Once conditions in equations 4 and 5 are satisfied, full reinitialization is performed, where all the positions, velocities, global best, …, etc. are restarted. This is a necessary procedure for most cases to capture the new global best.

However, for less complicated systems, where the shifted global best resides in the vicinity of the old one, a full reinitialization has a draw back on the efficiency of the system. Hence, as an intermediate step, a partial reinitialization, where only the global best is restarted, could be performed avoiding full reinitialization. Fig. 1 illustrate the proposed scheme.

2. PV Model

The single-diode model of a PV cell depicted in Fig. 2 is used for modeling and simulating the PV system. The PV model includes a current source in antiparallel to a diode, the leakage current and ohmic losses are considered using a parallel and series resistance, respectively. Equations (6-11) define the output of the PV model and its temperature and irradiance dependence, with the parameters defined in Table 1.

$$I = I_{PV} - I_o\left[exp\left(\frac{(V+IR_S)}{V_t A}\right) - 1\right] - \frac{(V+IR_S)}{R_p}, \tag{6}$$

$$V_t = \frac{N_S K T}{q}, \tag{7}$$

$$I_{PV} = I_{sc} + (T_m - T_r)\left(\frac{S_r}{1000}\right), \tag{8}$$

$$I_o = I_{RS}\left(\frac{T}{T_r}\right)^{\frac{3}{A}} exp\left(\frac{-qV_g}{AK}\left(\frac{1}{T} - \frac{1}{T_r}\right)\right), \tag{9}$$

$$I_{RS} = \frac{I_{sc}}{exp\left(\frac{qV_{OC}}{N_S AKT}\right) - 1}, \tag{10}$$

$$I_{sc} = \left(I_{sc_r} + \alpha(T - T_r)\right)\left(\frac{S_r}{S_{r_r}}\right) \tag{11}$$

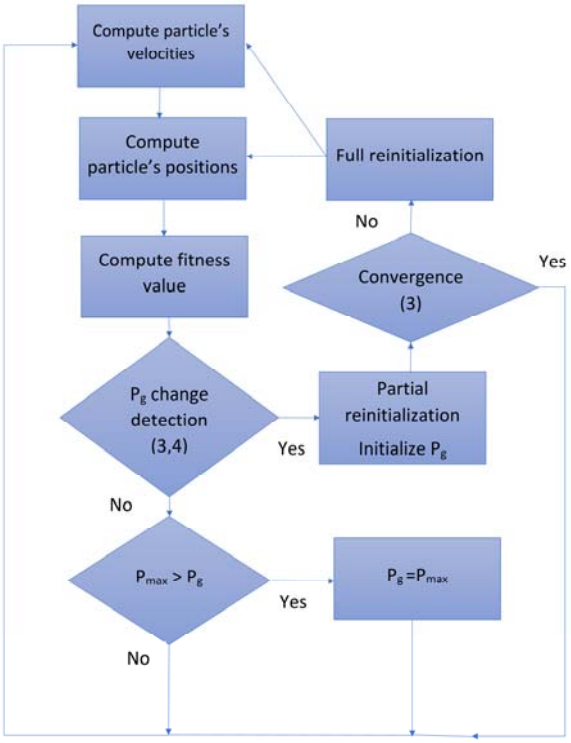


Fig. 1. The proposed reinitialization mechanism.

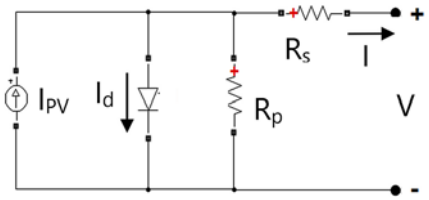


Fig. 2. Single-diode model of PV cell.

Table 1. Nomenclature of PV system.

I_{PV}	Photocurrent [A].
I	Photovoltaic Current [A].
V	Terminal voltage of PV module [V].
I_o	Diode Saturation current [A].
q	Charge of the electron [C].
K	Boltzmann's constant [J/K].
T	Junction temperature [K].
N_s	Number of cells in series.
R_s	Parallel resistance of PV module [Ω].
R_p	Series resistance of PV module [Ω].
A	Diode ideality constant.
I_{sc}	Short circuit current [A].
I_{RS}	Diode reverse saturation current [A].
S_r	Solar irradiance [W/m^2].
V_t	Thermal Voltage [V].
V_g	Band gap energy of semiconductor [eV].

Specifications for a single module are listed in Table 2. The I-V and P-V characteristics of the PV array that consists of 10 series modules and 40 parallel strings are shown in Fig. 3. These characteristics are plotted at 25 C° for different values of irradiance (1, 0.7 and 0.5 KW/m^2). It is clearly seen that the maximum values of power are captured at an irradiance of 1 KW/m^2 and 0.7 KW/m^2, corresponding to 290 V and 292.6 V, respectively. This behavior makes the system in hand a good candidate to test the proposed partial reinitialization technique.

Table 2. Specifications of a single PV module.

Max power (W)	213.15
Open circuit voltage (V)	36.3
Voltage at max. power (V)	29
Short circuit current (A)	7.84
Current at max. power (A)	7.61
Number of modules in series	10
Number of strings in parallel	40

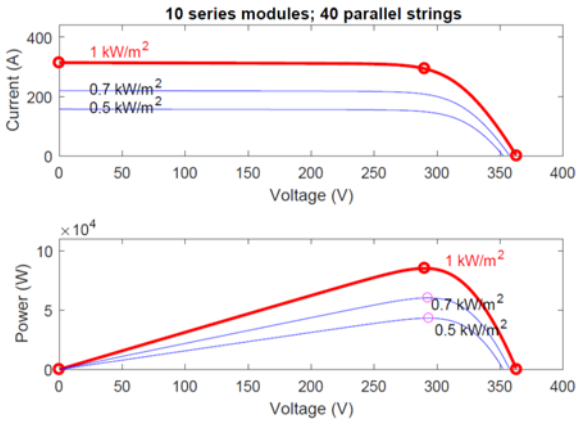


Fig. 3. I-V and P-V characteristics for different irradiance levels.

3. MPPT of PV Module via the Proposed PSO

The PSO-DIW [8] algorithm parameters (c_1, c_2, w, w_{damp} & N) used in this work are enumerated in Table 3. The proposed PSO is implemented on the PV module, where the voltage is manipulated based on power measurements. Fig. 4 and 5 shows the simulated PV module under variable irradiance; 1 KW/m^2 (0-0.22 s) and 0.7 KW/m^2 (0.22-0.45 s). It can be seen from the aforementioned figures that the response requires ~ 0.2 s to converge to MPP via full reinitialization. However, when the irradiance changes to 0.7 KW/m^2, it is detected by conditions in equations 4 and 5 and the partial reinitialization is activated. Fig. 4 depicts how the controller manipulates the voltage in response to partial reinitialization. The response settled within ~ 0.06 s from the activation, with much less fluctuations in the voltage and consequently in the power, compared to the initial response. The figure also shows that there is no need for a full reinitialization when the irradiance changed. Otherwise, as shown in Fig. 5, a full reinitialization will require more time to capture MPP, and it will reduce the output power during the long convergence period. These results show that partial reinitialization has better performance compared to full reinitialization is done, or even worse when no reinitialization is carried out as.

Table 3. Parameters of the PSO-DIW.

c_1	1.4962
c_2	1.4962
w	0.7298
w_{damp}	0.99
N	10

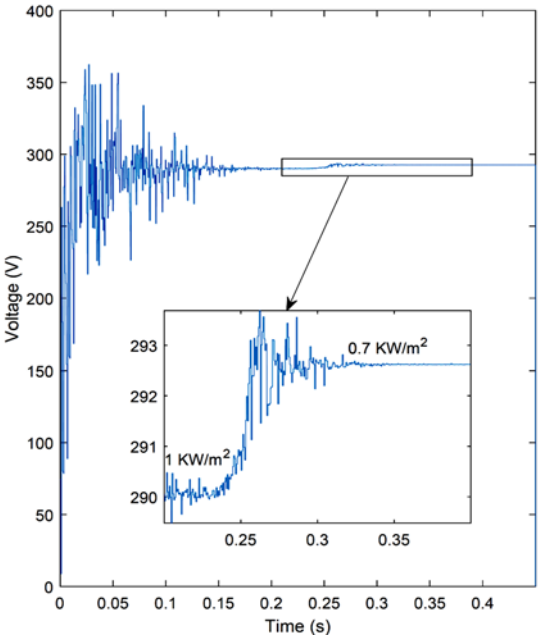


Fig. 4. MPPT using PSO with partial reinitialization.

4. Conclusions

The PSO optimization technique was successful in achieving online optimization of the performance of a PV system. The need for reinitialization of PSO is motivated by the variation of the MPP of the PV system due to varying operating conditions. The reinitialization technique proposed in the PSO algorithm was successful in maintaining a maximum power production under changing irradiance conditions without the need to a full reinitialization. Partial reinitialization requires less convergence time, much less fluctuations, and consequently capture more power, in the PV system case. Moreover, the proposed approach introduces more opportunities for online PSO based control technique, with its better performance and adaptivity in solving multidimensional, multiobjective, and constrained problems [8].

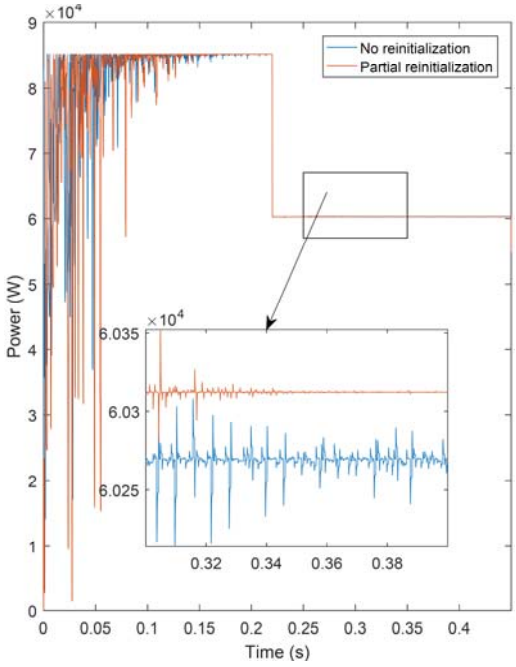


Fig. 5. MPPT using PSO with partial reinitialization and without reinitialization.

References

[1]. J. Kennedy, R. C. Eberhart, Particle swarm optimization, in *Proceedings of the IEEE International Conference on Neural Networks*, 1995, pp. 1942-1948.

[2]. S. Kiranyaz, T. Ince, M. Gabbouj. Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition, *Springer*, Berlin, 2014.

[3]. M. Miyatake, M. Veerachary, F. Toriumi, N. Fujii, H. Ko, Maximum power point tracking of multiple photovoltaic arrays: A PSO approach, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 47, Issue 1, 2011, pp. 367-380.

[4]. Y. Romasevich, V. Loveikin, A novel multi-epoch particle swarm optimization technique, *Cybernetics and Information Technologies*, Vol. 18, Issue 3, 2018, pp. 62-74.

[5]. A. Martínez-Cruz, R. Barrón-Fernández, H. Molina-Lozano, M. A. Ramírez-Salinas, L. A. Villa-Vargas, P. Cortés-Antonio, K. T. T. Cheng, An automatic functional coverage for digital systems through a binary particle swarm optimization algorithm with a reinitialization mechanism, *Journal of Electronic Testing*, Vol. 33, Issue 4, 2017, pp. 431-447.

[6]. N. K. Rout, D. P. Das, G. Panda, Particle swarm optimization based active noise control algorithm without secondary path identification, *IEEE Transactions on Instrumentation and Measurement,* Vol. 61, Issue 2, 2011, pp. 554-563.

[7]. A. M. Eltamaly, M. S. Al-Saud, A. G. Abo-Khalil, Performance improvement of PV systems' maximum power point tracker based on a scanning PSO particle strategy, *Sustainability*, Vol. 12, Issue 3, 2020, 1185.

[8]. B. Jiao, Z. Lian, X. Gu, A dynamic inertia weight particle swarm optimization algorithm, *Chaos Solutions & Fractals,* Vol. 37, Issue 3, 2008, pp. 698-705.

(025)

Important Data Quality Accents for Data Analytics and Decision Making

Ina Naydenova [1], Zlatinka Kovacheva [1, 2] and Kalinka Kaloyanova [3, 1]
[1] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Acad. Georgi Bonchev Str., Block 8, 1113 Sofia, Bulgaria
[2] University of Mining and Geology, Sofia, Bulgaria
[3] Sofia University "St. Kliment Ohridski", Faculty of Mathematics and Informatics
Tel.: + 359885193847
E-mail: naydenova@gmail.com

Summary: Data quality is a key factor for the successful data analytics and decision making in modern companies and organizations. Based on our long-term experience and consulting in a lot of business projects in various industrial sectors, we are focusing on the major factors for data quality provisioning which are usually poorly or insufficiently understood. In the present contribution, we are sharing a method for data unification that has been used successfully for over 15 years.

Keywords: Data quality, Data analytics, Decision making, Data validation, Data unification.

1. Introduction

The amount of data available to decision makers has risen exponentially in recent years. More people have access to more data, but few are increasing the rate at which they use data to make decisions. According to a recent survey, about two-thirds of business decision makers have some reservations or active mistrust in their data and analytics [1]. In addition, 36 percent of CEOs admit they cannot make data-driven decisions until they invest significantly in data quality [2]. So, in order to establish confidence in analytics, organizations need effective data quality (DQ) programs. Here, we will pay attention to some important data quality accents that are usually not fully understood or underestimated, and represent a significant setback in compiling and implementing a successful strategy for improving data quality. In addition, we will briefly describe how we solved the problem of duplication of company clients in various IT systems, which was of key importance for achieving the required data quality level for one of our industrial customers.

2. Poorly or Insufficiently Understood Aspects of Data Quality Provisioning

Based on our long-term experience and consulting in many business projects in various industrial sectors such as banking, telecommunication services, insurance, utilities and others, we can summarize the following issues as a main factor that should be taken into account during the data quality improvement initiatives:

Contextual nature of data quality

Many professionals are not aware of the contextual nature of data quality. The idea that some data is good ("high quality") for some purposes but is poor ("low quality") for others is not always well assimilated. Incorrect data may have a different impact on the business, depending on the purposes for which the data is used. For example, if a wrong first customer name is registered in an IT system, this indicates the presence of poor data. However, for the task of mail notifications, where only the customer's last name is used, there is no quality issue with the system' data. At the same time, the wrong first name could be a problem if the person's identity needs to be verified. Poor data is often associated with the discrepancies between reality and the data stored in the system, but the quality of data is a much broader concept. A similar observation was shared back in 1990 in [3]. The perception that data quality should be defined in accordance with the consumer's needs is further developed in [4]. There "data quality" is defined as data that is "fit for use" by data consumers. The authors also define a "data quality dimension" as a set of data quality attributes that represent a single aspect or construct of data quality. In subsequent studies scholars propose a variety of data quality dimensions. The majority of the authors include some versions of accuracy, completeness, consistency and timeliness among them [5]. It should be noted that quality dimensions give some structure to data quality expectations, but they have nothing to do with the task-dependent nature of the data quality. Let us take the completeness, for example: for the task of mail notification we need a complete and accurate mail address, but for the purposes of geographic segmentation only the city part of the address is required [6].

The contextual nature of the quality is a root of many challenges. The lack of understanding of the duality of data quality (objectivity and subjectivity) leads to a number of problems, such as inappropriate approach in setting data quality goals, difficulties in data quality issues assessment, incorrect expectations regarding the

DQ provisioning processes and the business value to be achieved. It also raises the question of the usability of solutions that accumulate large amounts of data from various sources, without clear vision about their purposes of use (e.g. Big Data solutions).

Task-dependent objective measurement

How good is a company's data quality? Answering this question requires usable data quality metrics. Companies must deal with both the subjective perceptions of the individuals involved with the data, and the objective measurements based on the data set in question. Subjective data quality assessments reflect the needs and experiences of stakeholders: the collectors, custodians, and consumers of data products. It should be borne in mind that the objective assessments could be task-independent, but could also be task-dependent as well. Task-independent metrics reflect states of the data without the contextual knowledge of the application and can be applied to any data set, regardless of the tasks at hand, while the task-dependent metrics are developed in specific application contexts [7]. According to our observations, most of the business significant data quality problems in modern organizations require the use of task-dependent measurement. Also, there is a common expectation that objective data quality improvement necessarily implies business value. Unfortunately, the limited awareness of what data quality improvement can truly imply often drives technical approaches that don't always translate into improving the business. Objective data quality metrics (such as a number of invalid values or percentage of missing data elements) may not necessarily be tied to the business's performance [8].
Many publications and training courses discuss various assessment techniques and metrics such as scaling and normalization, different types of counters and ratios, error rates, distribution and distance functions, weighted metrics, standard deviation, etc. Different groups of metrics are usually discussed as appropriate for different quality dimensions. However, guidelines and good practices of how to relate and adapt the objective metrics to the business context and consumer expectations is rarely discussed. This trend is particularly pronounced in publications related to quality assurance in big data solutions. Thereby the impression that the quality of data and its measurement is an objective one and should be done through appropriate IT tools by the IT specialists in the organization is reinforced.

The importance of data quality requirements analysis

Another common mistake is to underestimate the preliminary analysis to identify the extent of data quality issues and to draw up the quality improvement strategy. The allocated time and resources for the analysis phase usually are not enough. Many business professionals are not aware that they have to become part of such analysis, so they are not interested in the project aims. Other participants in turn focus on analyzing and profiling the data quality state in the databases, instead of interviewing the business users about the problems encountered by them. The organization remains unable to see the big picture and it is not able to decide how to achieve maximum benefit with minimum effort. The methodologies for data quality improvement also reflect in the underestimation of the need of a holistic analysis. Almost all methodologies include activities which examine data schemes and perform interviews to reach a complete understanding of data, but only a few of them consider the data quality requirements analysis step, identifying data quality issues from the business perspective and collecting new target quality levels [9].

Whose responsibility is to provide quality data

One of the main factors for the success of data quality initiatives is to understand correctly whose responsibility it is to provide quality data. According to our experience, the misconception that data quality is mainly an IT problem and it could be solved by purchasing sophisticated (and often expensive) data quality tools is very common. One of the reasons for such a delusion is the assumption that developers bear primary responsibility for ensuring that the software they write works properly, no matter which data is fed into it. Since they write the code, they have the power to control how an application will respond when it receives low quality data. In the real world even the best programmers cannot foresee every type of data quality issue that could occur within their applications. Even if they did, the applications might end up being overloaded by functions that handle obscure data quality issues [10]. Actually, the organization and its processes are always more important than the technology since they are defined according to company strategy [11]. The business professionals must not only be involved, but also lead the data quality initiatives. This is a consequence of the contextual nature of the quality. In fact, the businesses know the context and the indented uses of data better than anyone does.

Given the strategic importance that data plays in competitive markets, high data quality can only be achieved if IT and business find a mutually intelligible language, a mode of operation, that overcomes the traditional disconnect that can be observed in so many corporations. Conceptually what needs to happen to enable business alignment is bringing together problem holder and problem owner. The problem holder is the person who experiences "pain" from a problem; a problem owner is the person who controls the resources needed to resolve it. The problems resolutions lie in making the problem owner feel some of the problem holder's "pain" so he becomes motivated to do something about the problem [12].

Systems interoperation in the modern business

The lack of overall vision for the data flows between information systems of the organization and the lack of effective strategy for synchronizing the changes in them often lead to the presence of inconsistent, obsolete and contradictory data. The managers rely on the corporate data warehouses as a "single source of truth". However, in modern data warehouses, which accumulate a huge amount of data from different operating systems, this is difficult to achieve in practice. Moreover, the operating systems that feed data warehouses are also used for an extraction of various reports. Correction of bad data must be done in all systems where copies are stored. Correcting them only in the data warehouse will improve the quality of data locally but will enlarge another severe problem – the existence of conflicting information at the organizational level.

Identical objects recognition

A serious problem in some organizations is the resolution of cases with duplication of data, when there are no clear criteria by which it is possible to identify two objects as identical (e.g. if some individuals do not have the same national identification number in different data sources). Similar problems are particularly relevant when collecting information from various sources and information channels in large data warehouses, data lakes and big data solutions. To solve such problems, record linkage techniques are recommended in the literature. Data linkage can be deterministic (exact), where each matching variable needs to agree exactly to determine a correct match, or probabilistic (fuzzy matching), where there are data errors in the matching variables and a probability of a correct match should be determined [13].

3. A Method for Data Unification

In the previous section, we outlined some key points that we consider important enough to be taken into account by organizations looking to improve the quality of data they work with. In this section, we will share a specific example of a successful implementation, which has been used productively for more than 15 years. We will briefly present the method for a unification of individuals and legal entities records, which belongs to the class of probabilistic record linkage techniques. The method merges customers' information and eliminates duplicated objects across several operational systems. During the implementation, we had to use heuristics that take into account the context and specifics of the data quality issues of the sponsoring organization – a holding company with different types of businesses.

The method we offered defines numerical distance estimation between similar objects and uses a neighborhood based approach to identified clusters of identical customers. It has three main stages: data preparation and preprocessing stage, identical candidates pairs collection and candidates evaluation.

The first stage involves activities such as splitting of unformatted data and fields preprocessing. The goal of this stage is to derive new attributes values that will be used in the next two stages. First, we do some transformations. For example, we eliminate common and therefore less meaningful words that could bring an unnecessary noise in data for particular string attributes. Different transformation rules are applied to key individuals and legal entities characteristics such as names, mail addresses, individuals' identity documents, legal entity registration entries and others. Then we do some validations based on the statistical information derived from the source data. The goal of the validations is to eliminate artificial and non-identifying attribute values, but to preserve object specific values even if they are of dubious fidelity. Insufficiently informative incorrect data are removed.

During the second stage, the available data is ransacked for very likely identical objects. The goal is to form a reduced list of pairs which will be precisely estimated in the next stage. The pairs of objects that do not have a chance to pass the identity threshold are ignored as uninteresting for further evaluation.

The goal of the last stage is to calculate the similarity measure of every pair in our candidates lists as a weighted sum of similarity degrees between attributes. We give specific positive scores when the values of a particular attribute are equal for a candidate pair. The scores' magnitude is defined on the base of a heuristic grade of the attribute importance related to the objects identity. We also give a penalty (negative scores) when some key attributes are different. Since it is possible for a customer A to be identical with B, and B to be identical with C, after the pair evaluation we apply an algorithm to discover the transitive closure of the identity relation.

As can be expected, after the initial implementation of the method we went through a period of weights adjustment and elimination of problems arising from specific business practices that we were not aware of during the method implementation. For example, we found that during the preparation of an insurance contract, the insurance agent enters his own personal data occasionally. These are situations where the agent does not have the customer's data, but is in a hurry to make a deal. Later he corrects the data with the customer's actual ones. In these cases, the unification algorithm brings together all the customers that are served by the same agent and unites them with the identical customers loaded by the other data sources. This practice, as well as some other discoveries in the data, provoked appropriate algorithmic amendments for the achievement of higher stability and better accuracy of the method [14].

4. DQ Implementation Considerations

In accordance with the data quality highlights discussed above, we will outline some implementation considerations and guidelines.

Relate data quality goals with business risks

When defining the objectives, scope and implementation strategy of the DQ program, do not forget to be guided by the contextual nature of the quality. A company has to measure data quality from a "risk" perspective in line with its business goals. The efforts required to reach an acceptable level of data quality must be proportional to the risk exposure. For instance, it is useless and quite costly to maintain mailing addresses up to date if the company uses them for purposes where the timeliness of the data is not essential. It's up to the "business", and according to the importance of each of its missions, to define and enforce a certain level of data quality requirement, since it's understood that data quality cannot be 100 % certified. This risk gets bigger when the data are used to feed the decision processes as part of Business Intelligence or are traded with external partners to the company (customer, supplier, regulatory authorities, etc.) [15]. The quality of data affects so many aspects of a company's business that bad data is even identified as a risk factor for the enforcement of the organization's security standards [16] and regulatory compliance (sensitive data may be missed and not encrypted due to data errors, security alerts could not be triggered properly, non-unified customer information can lead to incorrect anti-money laundering regulatory reporting, etc.) [17]. Of course, this is easier to be said than to put it into practice. Accurate measurement of the degree of risk and the return on investment of solving a data quality problem is often more challenging than it seems. There are many factors that determine the DQ initiatives cost and most of them are difficult to be measured (e.g. equipment and training costs, reputational costs, customer dissatisfaction costs, maintenance costs, excess labor costs, time costs of viewing irrelevant information, lost and missed opportunity costs, etc.). However, the mission is not impossible – some techniques such as focusing on particular issues and the use of subjective evaluation metrics can be useful and lead to very satisfactory results.

Focus your efforts on particular activities

The definition of the DQ goals should start with the identification of data related to the critical activity of the business and the efforts should be directed towards constant improvement of data quality for that particular activity. As a first step, the business's view of the existing data quality problems in the organization should be examined – which of the problems they identify as a high priority and which as secondary, what is the expected business value of eliminating the quality issues, etc. This survey could be done by means of a series of interviews and questionnaires. In the next step, the primary problems need to be analyzed in greater detail, both from a business and a technical point of view. As a result of this analysis, objective and/or subjective metrics can be defined to measure the problem. It should be borne in mind that object metrics are often task-dependent and must be customized to the context of specific problems. For example, when the business experts alert that there is an issue with the completeness of customer insurance income data, the IT specialist can easily calculate a ratio between rows with missing and non-missing income values. However, the business issue may only concern customers who have signed a certain type of contract, missing income could be a normal situation for customers who are foreigners, a dummy "zero value" could be entered in certain situations, etc. These specifics must be reflected in the data quality metric.

Do not be afraid to use subjective metrics

In cases where it is difficult to define an objective metric for assessment of the severity or the cost of DQ issues, the estimation should be based on the subjective judgments of the business experts. For example, regarding the question of how bad data affects the organization's reputation with customers, we can launch a survey to appropriate people in the organization who have direct observations on the issue. The tricky point here is how to determine the right people to whom the questionnaires should be addressed, as well as choosing the appropriate format through which to conduct the survey. In addition to a quantitative assessment to be given based on people's subjective experience, it is good practice to ask for a brief explanation or examples illustrating the experience on the basis of which the assessment was given.

Compromise with the accuracy of the assessment

Be careful not to make the "formula" for evaluating a DQ issue too complex (more important is to be clear). With the accuracy of the assessment, it is often necessary to make compromises. It is better to ignore some insignificant cost factors than to take the risk of getting lost in the assessment complexity or to present the assessment as a black box to the management. If you feel uncertain about the significance of certain cost factors, then you can resort again to a survey to check the opinion of the business professionals about the degree of impact of the factors based on their subjective experience. If there are serious discrepancies in the assessment, additional analysis can be conducted - more detailed interviews with the "business" or efforts to define objective metrics for measurement. Alternatively, ignoring cost factors could be based on the assumption that they are embedded in the subjective opinion of data consumers to some extent. In general, business experts intuitively prioritize as more severe the problems that lead to greater losses for the business. However, this is not always entirely true. For example, people tend to classify the problems that generate much workload and complicate business processes as more severe. Simplifying the process does not always lead to greater direct profits, but it certainly has an indirect effect, saving human labor.

Do not spare effort and resources for the analysis of the current state

Allocate sufficient time and resources for the analysis of the available data quality issues in the organization, along with the definition of the DQ program objectives. The outcomes of the "as-is" analysis (or assessment phase according to some methodologies) are of key importance for the success of the program. Ensure the commitment of both "Business" and IT departments. Make sure the communication between them is at a good level. This will allow the objective assessment to be properly adapted to the business context and a suitable improvement strategy to be designed. Due to the importance of the outcomes of the analysis, it is advisable to be launched as a separate project.

At a minimum, the assessment phase should include the following activities [18]:

- *DQ requirements analysis* to identify quality issues and set the program goals and new quality targets;
- *identification of critical areas*, which selects the most relevant databases and data flows to be assessed quantitatively;
- *data analysis*, which examines data schemes and performs interviews to reach a complete understanding of data and related architectural and management rules;
- *process modeling*, which provides a model of the processes producing or updating data;
- *measurement of quality*, which selects the quality dimensions affected by the quality issues identified in the DQ requirements analysis step and defines corresponding metrics.

The analysis should also include a study of the intersystem data connectivity. Due to the technical and implementation complexity, this aspect is often neglected. Many DQ initiatives are launched to improve the quality of data in a particular application, but nowadays the information handled by the decision makers and senior management levels is global and integrated. It is received via several channels and the resulting discrepancies lead to mistrust in data and their reliability.

Once the analysis of the "as-is" state is done and the program objectives are presented, the definition of the implementation strategy could start. Typical activities of this phase are the assignment of the processes' and data' responsibilities, selection of improvement approaches, techniques and tools, design of improvement solutions, including redesign of business processes, data workflows, validation procedures, etc. The definition of strategies for management and monitoring of the improvements is also a part of this stage [18].

According to [19] there are six aspects of an organization's operations that have to be taken into account when creating your data quality strategy, namely: (1) *Context* – the data being cleansed and the purposes for which it is used; (2) *Storage* – where the data resides, including technologies, platforms and system architecture aspects; (3) *Data Flow* – how the data enters and moves through the organization; (4) *Work Flow* – how work activities interact with and use the data; (5) *Stewardship* – people responsible for managing the data; (6) *Continuous Monitoring* – processes for regularly validating the data. Each of these aspects or perspectives of the operational environment required different actions, resources and decisions to be considered. Where the data quality strategy formulation process ends, data quality project management takes over. While the goals have been documented, and the data sets established, the project manager must build the project requirements from the goals. The project manager should adhere to sound project management principles and concepts that apply to any project, such as task formulation, estimation, resource assignments, scheduling, risk analysis, mitigation, and project monitoring against critical success factors. Few of these tactical issues are covered in a strategy-level plan.

Do not underestimate the human factor and working collaboration

No strategy is complete without the evaluation of the human factor and its effect on operations. Work flows and data flows are initiated by people. Carefully define the participants in the program, their roles and responsibilities. On the base of the provided analysis of the DQ issues consider who are the problem owners and who are the problem holders. Inform them of the plans, ask each one about their specific needs, and collect their feedback. If there are many stakeholders, selecting a representative from each user function is highly encouraged [19]. Consider the impact of new processes or changed data sets on organizational structure and make sure people don't react negatively to these changes for fear of losing the achieved current conditions. Also, do not forget to plan appropriate training for updated definitions, new processes and model changes in order to support better user acceptance and proper future maintenance.

Keep in mind that maintaining good data quality is an ongoing process

Providing data quality should not be considered as a one-time initiative. It is an ongoing process. The one-time initiatives have their place, but DQ problems that can be permanently fixed by a single action are rare. For DQ issues which require a process-driven improvement strategy, be sure to secure the data quality maintenance strategy too.

5. Conclusions

Nowadays, almost every company struggles against data quality issues to some extent. In the conditions of modern globalization and integration of information particularly relevant is the problem with the contradictory data, which leads to a serious lack of

trust in the performed analyses. Many people in modern companies spend a lot of their working time validating the results of different types of reporting, clarifying the causes of inconsistencies in them, testing and comparing data. Business analysts lose motivation to perform advanced analyses and senior managers rely more on their intuition in defining the company's strategy, instead of making a data-driven decision [20]. In today's information age, this is a rather surprising phenomenon.

Data quality has been a topic of interest to the scientific community for more than 30 years. Accumulated knowledge and experience are significant, but the topicality of the DQ issues shows that there is still much to be done. Our experience shows that most organizations have the same class of problems and obstacles to achieving better data quality. Here, we have paid attention to only a few of them, which we consider as the most common and important for effective data analytics and decision making, but their list is much longer for open discussion. In our opinion, efforts should be focused not only on the development of new techniques, sophisticated data quality tools and methodologies, but also on the promotion of common mistakes, as well as the sharing of good practices.

Acknowledgement

This research is supported by the project BG05M2OP001-1.001-0004 Universities for Science, Informatics and Technologies in the e-Society (UNITe) and the National Scientific Program 'Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)'.

References

[1]. B. Violino, Many executives lack a high level of trust in their organization's data, analytics, and AI, Red Ventures Znet, https://www.zdnet.com/article/most-executives-dont-trust-their-organizations-data-analytics-and-ai/

[2]. KPMG, Trusted analytics matter to CEOs, Trusted Analytics Article Series, https://assets.kpmg/content/dam/kpmg/xx/pdf/2017/08/da-trusted-analytics-ceo-outlook.pdf

[3]. R. Wang, L. Guarascio, Dimensions of Data Quality: Toward Quality Data by Design, IFSRC Discussion, Paper #CIS-91-06, https://web.mit.edu/smadnick/www/wp2/1991-06.pdf

[4]. R. Wang, D. Strong, Beyond accuracy: What data quality means to data consumers, *Journal of Management Information Systems*, Vol. 12, Issue 4, 1996, pp. 5-34.

[5]. L. Sebastian-Coleman, Measuring Data Quality for Ongoing Improvement – A Data Quality Assessment Framework, *Morgan Kaufmann*, 2013.

[6]. J. Todorov, I. Sturbanova, M. Trifonova, Information system for planning, management and reporting of open cast mines production (output), in *Proceedings of the First International Conference on Information Systems & Datagrid*, Sofia, 17-18 Feb. 2005, pp. 147-154.

[7]. L. Pipino, Y. Lee, R. Wang. Data Quality Assessment, Communications of the ACM, July 2013, http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCACMApr02.pdf

[8]. D. Loshin, The Practitioner's Guide to Data Quality Improvement, *Morgan Kaufmann*, 2011.

[9]. C. Batini, C. Cappiello, C. Francalanci, A. Maurino, Methodologies for data quality assessment and improvement, *ACM Computing Surveys*, Vol. 41, Issue 3, 16.

[10]. Ch. Tozzi, Solving Data Quality Problems Is Not (Only) Programmers' Responsibility, 2020, Precisely, https://www.precisely.com/blog/data-quality/solving-data-quality-problems

[11]. BARC, Data Quality and Master Data Management: How to Improve Your Data Quality, 2020, Business Application Research Center, https://bi-survey.com/data-quality-master-data-management

[12]. T. Breur, Data quality is everyone's business – Managing information quality – Part 2, *Journal of Direct*, *Data and Digital Marketing Practice*, Vol. 11, 2009, pp.114-123.

[13]. N. Shlomo, Overview of data linkage methods for policy design and evaluation, in Data-Driven Policy Impact Evaluation (N. Crato, P. Paruolo, Eds.), *Springer,* Cham, 2019.

[14]. I. Naydenova, K. Kaloyanova, S. Ivanov, Multi-source customer identification, in *Proceedings of the Third International Conference on Information Systems & Grid Technologies*, 28-29 May 2009, Sofia, Bulgaria, pp. 77-85.

[15]. Rever Data Engineers, Why data accuracy is a main issue for companies, 2020, Rever SA, https://www.dataengineers.eu/en/why-data-accuracy-is-a-main-issue-for-companies/

[16]. ISO, Information technology – Security techniques – Information security management systems – Requirements, ISO/IEC 27001:2013, https://www.iso.org/standard/54534.html

[17]. Informatica, Data Quality: The Hidden Enabler of Security and Compliance, 2018, Integrationworx Ltd, http://www.integrationworx.net/sites/default/files/downloads/dataqualitythehiddenenablerofsecurityandcompliance.pdf

[18]. C. Batini, C. Cappiello, C. Francalanci, A. Maurino, Methodologies for data quality assessment and improvement, *ACM Computing Surveys*, Vol. 41, Issue 3, 2009, 16.

[19]. F. Davis, Data Quality Strategy: A Step-By-Step Approach, in *Proceedings of the Ninth International Conference on Information Quality*, 2004, http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%202004/Papers/DQStrategy.pdf

[20]. B. Fisher, Five ways CEOs can build a data-driven organization, Morgan Kaufmann, 2019, KPMG, https://home.kpmg/xx/en/home/insights/2019/04/dont-doubt-the-data.html

(026)

Internet Video Traffic Classification with Convolutional Neural Networks

E. Grabs [1], E. Petersons [1], D. Efrosinin [2], A. Ipatovs [1] and V. Sturm [2]
[1] Riga Technical University, Telecommunications Institute, Azenes str. 12, LV-1048 Riga, Latvia
[2] Johannes Kepler University, Institute of Stochastics, Altenberger Strasse, 4040 Linz, Austria
E-mail: elans.grabs@rtu.lv

Summary: The presented article provides results on Deep Learning model application for internet traffic classification. The classification task is performed for video-traffic of two types: real-time streaming video and on-demand video record. Multiple Machine Learning methods have been used to solve the specified task, and this article presents results obtained by Convolutional Neural Networks method. The article describes data pre-processing and formatting, as well as used model parameters. The article is concluded with some of the obtained results presented in form of Tables and 3D surface plots. The conclusion finalizes the article.

Keywords: Deep learning, Convolutional Neural Networks, Computer network traffic, Traffic stochastic features.

1. Introduction

During the last few years, the fast development of big data and Machine Learning applications has been observed. These topics are currently widely studied in many disciplines, and computer networks traffic is no exception. There are many tasks in computer networks, which can be solved by different methods of Machine Learning [1].

One of such task groups are classification tasks, more specifically: the classification of different video-traffic modes [2]. For this purpose many Machine Learning algorithms can be used, but in the presented results the Convolutional Neural Networks (CNN) have been used, which is one of Deep Learning methods. These methods are also being used for computer network traffic [3, 4]. The similar task has been performed for classification of WEB, Audio and Video traffic in [5]. In the presented by authors of this abstract research, however, multiple video traffic types are classified, in contrast to Video/Non-video traffic classification in references above.

2. Data Collection and Preparation

For network traffic capturing the WireShark software has been used. This software is capable of capturing all packets sent to and from specified network interface. However, only statistics of the packet arrival time and size have been exported in CSV format file.

This flow of packets has been further sliced into segments with sampling time interval T, which for current experiments has been set to either of following three values: 1 ms, 10 ms or 100 ms. After this operation, the traffic is stored in a form of packet intensity values over specified time interval.

This traffic intensity has been further analyzed by a window of specific length (from 0.1 sec to 10 sec), and for entire window the statistical moments have been calculated. These statistical moments (a total of 6 values) have been further used as features of analyzed traffic.

The procedure has been performed for both traffic types (Live and Record), and after that, the samples of features have been randomly mixed and marked.

The resulting data formed a matrix, where each contains 6 feature values for each segment of traffic intensity, processed by a window of specified length.

3. Convolutional Neural Network Model

The training data has been generated by randomly mixing two types of network traffic with the same ratio, whereas all unused traffic has been used to create a testing data set in a similar way.

The window processed data is afterwards loaded in Python, where Keras with TensorFlow is used to create a CNN Deep Learning model. The structure of used model is still being studied, but for presented results the following model architecture has been used:

1. 1D Convolution layer with 10 filters and 2 to 6 kernel length, ReLU activation function;
2. Flatten layer without additional parameters;
3. Dense layer with 2 outputs and SoftMax activation function.

For calculations of neural network weights the Adam optimizer has been used, for a total number of 25 epochs (runs). Due to random nature of optimizer initial conditions, several runs (5) have been performed to make sure, the accuracy has been evaluated correctly, and the highest accuracy result has been selected.

4. Analysis of the Results

The research has been made for different traffic and model parameters: traffic intensity time interval, window length, length of CNN kernel, number of statistical moments, etc.

For example, Table 1 shows some accuracy results for different traffic parameters (intensity and window length), whereas Fig. 1 displays 3D surface plot of accuracy depending on window length and CNN model kernel length.

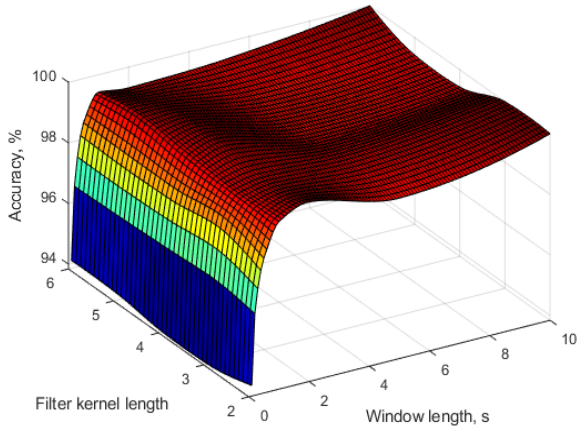


Fig. 1. 3D surface plot for RAW 1080p60 HD Quality Video classifier results with different window size and kernel length.

Table 1. Deep Learning accuracy evaluation for different traffic analysis parameters (1080p60 HD video quality mode).

Traffic window length	Accuracy for different traffic intensity time interval		
	100 ms	10 ms	1 ms
0.1 sec	94.22 %	94.10 %	94.30 %
0.5 sec	99.17 %	98.59 %	96.39 %
1 sec	99.48 %	99.08 %	97.82 %
5 sec	99.71 %	99.14 %	99.71 %

5. Conclusion

In most cases of the model training and accuracy estimation, the accuracy of traffic classification is high enough (up to 99 %, as is shown in Table 1).

Another important conclusion to be made is the fact, that it is not necessary to increase Filter kernel length, this value can be set as low as 2, in most cases. Although, slight increase may actually lead to accuracy improvement.

Concerning the window length of the traffic analysis, its value has the highest influence on classification accuracy. However, caution must be used, since if window length is too high, the traffic may change over that time duration.

Acknowledgements

This work has been supported by the European Regional Development Fund within the Activity 1.1.1.2 "Post-doctoral Research Aid" of the Specific Aid Objective 1.1.1 "To increase the research and innovative capacity of scientific institutions of Latvia and the ability to attract external financing, investing in human resources and infrastructure" of the Operational Programme "Growth and Employment" (No.1.1.1.2/VIAA/2/18/332).

References

[1]. R. Boutaba, M. A. Salahuddin, N. Limam, et al., A comprehensive survey on machine learning for networking: evolution, applications and research opportunities, *J. Internet Serv. Appl.*, 2018, Vol. 9, 16.

[2]. M. Torres Vega, D. C. Mocanu, A. Liotta, Unsupervised deep learning for real-time assessment of video streaming services, *Multimed. Tools Appl.*, Vol. 76, 2017, pp. 22303–22327.

[3]. M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, J. Lloret, Network traffic classifier with Convolutional and Recurrent Neural Networks for Internet of Things, *IEEE Access*, Vol. 5, 2017, pp. 18042-18050.

[4]. H. Lim, J. Kim, J. Heo, K. Kim, Y. Hong, Y. Han, Packet-based network traffic classification using Deep Learning, in *Proceedings of the International Conference on Artificial Intelligence in Information and Communication (ICAIIC'19)*, Okinawa, Japan, 2019, pp. 046-051.

[5]. S. Kaoprakhon, V. Visoottiviseth, Classification of audio and video traffic over HTTP protocol, in *Proceedings of the 9th International Symposium on Communications and Information Technology (ISCIT'09)*, Icheon, 2009, pp. 1534-1539.

(029)

Human-Robot Interaction: Applications

Abdel-Nasser Sharkawy
Mechatronics Engineering, Mechanical Engineering Department
Faculty of Engineering, South Valley University
Qena 83523, Egypt
E-mail: eng.abdelnassersharkawy@gmail.com

Summary: Recently, human-robot interaction (HRI) is an extensive research topic and theme which gained importance and significance. HRI aims at the complementary combination between the robot capabilities and human skills. The robots assist humans in terms of precision, speed, and force. The humans contribute in terms of the experience, knowledge of executing the task, intuition, and easy adaptation and learning, and understanding of control strategies. In this work, the applications of human-robot interaction are reviewed. These applications could be industrial, medical, agricultural, servical, and educational. HRI can be found in industrial applications in picking and placing in the production lines, welding processes, parts assembly, and painting. Assistive robotics are one from the highest profile areas in HRI. For people with the physical and the mental challenges, the robots can provide the opportunity of interaction and therapy. In addition, HRI can be widely applied in hospitals. Nowadays, HRI is very important for facing the new coronavirus (COVID-19) pandemic. In agriculture, the cooperation between human and the robot helps with many tasks including harvesting, seeding, fertilizing, spraying, weed detection, hauling, and mowing. HRI can also be found in other applications such as education, mining, and home use.

Keywords: Human-robot interaction, Industrial applications, Medical and rehabilitation applications, Precision agriculture.

1. Introduction

HRI is the field of the study referring to understanding, designing, and evaluating the robotic systems used by or collaborating with the human operator. Interaction needs the communication between the robots and the humans. The human communication with the robot could take many forms. However, these forms are highly affected by whether the human operator closeness to the robot. Therefore, the communication or interaction between both the human and the robot is separated into two main categories [1, 2]: The first category is the remote interaction and the second is proximate interaction. In the remote one, the human does not exist in nearby the robot. In addition, they are spatially or even temporarily separated. An example for this category is the Mars Rovers which is separated from earth in space and time as well. In the proximate category, the human and robot coexist in the same location. An example for this category is the service robots during their sharing with the humans in the same room.

These categories can help distinguish between the applications requiring the mobility, physical interaction, or social interaction. Remote interaction using mobile robots refers to the tele-operation or the supervised control. Remote interaction using the physical manipulator refers to the tele-manipulation. Proximate interaction using the mobile robots takes the form of the robot assistant. Proximate interaction includes the physical interaction. In social interaction, the social and the emotive as well as the cognitive interaction aspects are included. In social interaction, the humans and the robots interact using the form of peers or companions. Importantly, social interactions with the robots take the form of proximate interaction rather than remote interaction.

From these categories, this manuscript addresses the proximate interaction. It presents some real applications for HRI in industrial environment, medical and rehabilitation, agriculture, education, and other environments. These applications show the importance of the interaction happening between the robot and human operator in our real life.

The rest of this paper is divided as follows. Section 2 presents the industrial applications of HRI. In Section 3, the HRI in the medical and rehabilitation applications is presented. HRI in agriculture is presented in Section 4, whereas Section 5 illustrates some other applications. Section 6 concludes the main important point of this manuscript.

2. HRI in Industry

HRI is widely used in industrial applications such as picking and placing in the production lines, welding processes, parts assembly, painting, and so on.

Some examples for HRI in the industrial application are presented from Figs. 1-4. The robot workstation is running in the plant of BMW in South Carolina in which the robot helps human operators to perform the assembly of the final door [3] (see Fig. 1). In the door assembly operation, human operator and robots work together. BMW plant has succeeded to implement and develop the direct human-robot cooperation and interaction in the series production.

Human-robot teams are found also in the flexible production lines [4, 5], as presented in Fig. 2. In this figure, the robotic manipulators and the human operators collaborate in handling of the work pieces. Safety is also is very important in such environments, since the proximity of the human operator to the robot that can lead to potential injuries.

Fig. 1. The robot helps the human performing the final door assembly in the BMW plant [3].

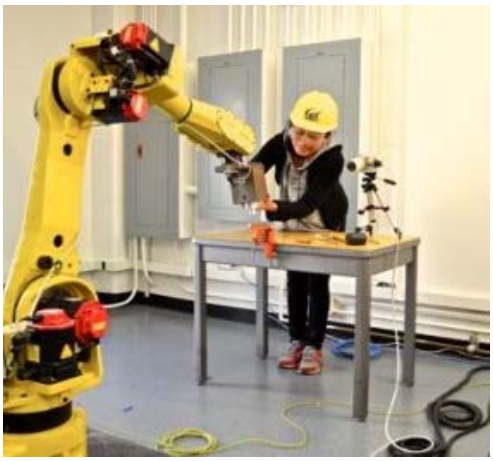


Fig. 2. The robot and the human workers cooperate in handling workpieces [4, 5].

Repetitive co-manipulation tasks as shown in Fig. 3 [6] can be performed in suitable poses of the human body. The poses can minimise the effects of the overloading joint torque. Furthermore, they can maximise the capacity of the manipulation of the human.

For handling of heavy and bulky components in welding situations, the multi-robot system with collaborative functionality assists the worker [7], as shown in Fig. 4. Two robots help position the components to be joined in the welding process, at which point the human operator carries out the welding task under favorable ergonomic conditions. In comparison with the standard welding bench, the human operator does not need to assume the uncomfortable postures or the work overhead. All necessary positioning and orientation of the work pieces can be performed by the robot. This also includes presenting of the components in the optimal position for the process of welding, allowing proper flow of the welding bead. Since the robotic repositioning motion is quite fast, the handling time which is about one third of the total process time is reduced to a minimum in comparison to welding processes in which HRI is not implemented.

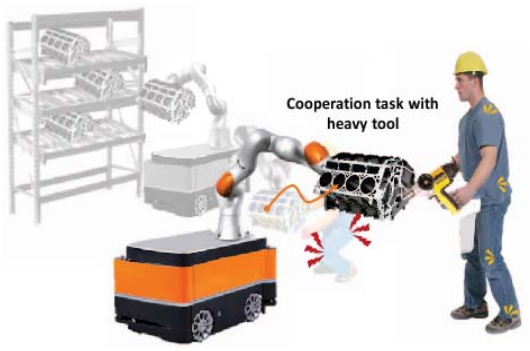


Fig. 3. Example for the repetitive co-manipulation tasks [6].

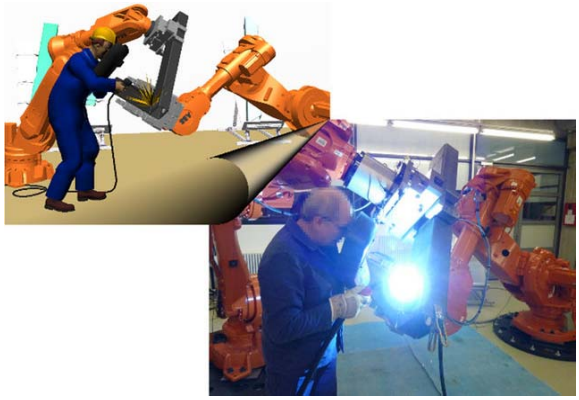


Fig. 4. The multi-robot system assists the worker in the welding process [7].

3. HRI in Rehabilitation and Medical Applications

Assistive robotics are one of the highest profile areas in HRI. For people with the physical and the mental challenges, robots can provide the opportunity for the interaction and therapy. This work is explored with autistic children in ref. [8, 9]. Many of them cannot respond strongly to the social cues but they respond very well to the mechanical devices. Robots give the possible therapeutic role for using the mechanical device for improving the social interactions [10]. Robots are being considered also for different domains in which children benefit, for example, children with experienced trauma. The social dimension of human-robot interaction is necessary not only in the assistive roles, but also in many areas and domains of the proximate interaction [11-15].

For people with physical disabilities, the robot embodiment provides the unique opportunities which are not possible with other technology forms. For example, the researchers are working on the design of the robots which provide and give the support for the physical therapy. Efforts include the providing of the prescribed force and the movement trajectories to help rebuild flexibility as well as strength [16]. Other work for detecting the motivational state and adjusting the therapy in order to maximize the benefits is presented in ref. [17]. Intelligent wheelchairs are the type of the robot which use the external sensors for supporting the

path planning and the collision avoidance for the person that requires the wheelchair [18, 19].

An example of the rehabilitation robotics was presented in [20], in which, a robotic system was built as shown in Fig. 5, that performs correctly the exercise for the conventional physiotherapy such as the shoulder flexion, in similar way to what the physiotherapist do during the co-manipulating with KUKA LWR robot. In this way, the excellent characteristics of the human (the capacity of the decision) and the robot (the precision, the work capacity, and the repeatability) could be combined to achieve better results in the musculoskeletal rehabilitation of the arm. Furthermore, the second as well as the main objective was evaluating the non-pathological shoulder behaviour while performing the shoulder flexion movements.

Fig. 5. An Example illustrates the rehabilitation robotics [20].

Coronavirus disease (COVID-19) is the infectious disease which is caused by the newly discovered coronavirus. The severity of COVID-19 symptoms can range from mild to severe. The cooperation between the human and the robot help to fight against the new pandemic. Two cases of the human-robot cooperation can be used in the hospital to fight against Coronavirus. The robot can add the solution to the nasopharyngeal swabs from the patients for detecting the coronavirus genetic material [21]. The positive test result illustrates that the patient is infected with the coronavirus. In that case, the samples which are taken from the patients' mouth, nose, and throat are tested for the genetic material of the coronavirus in the laboratory. The staff of the laboratory must only load the samples into the tray and then the COVID-19 test robot takes care with the pipetting. In the second case, a mobile unit is provided to the manipulator. The mobile robot works hand in hand with the human and align to the work piece with high precision. The mobile robot can be used for providing the patients with the food and medicine. In addition, the robot can measure the temperature of the patients. This would minimize the direct contact between the patients and medical staff and others, and therefore, minimizes the infection potential. The robot can also help the humans for sweeping and washing the floors and the walls. This also minimizes the virus infection.

4. HRI in Agriculture

The cooperation between human and robot in agriculture helps with many tasks such as harvesting, seeding, fertilizing, spraying, weed detection, hauling, and mowing.

In the precision agriculture as presented in Fig. 6 [22], the robot helps the human-operator or farmer in picking the strawberry. In that case, safety between the human operator and the robot must be included. Furthermore, the robot was controlled remotely by a co-located operator. Their task was to navigate the robot to the location of pickers when they requested it, allow the filled crates to be loaded onto the robot, and then transport these to the storage facility. In Fig. 7 [23], intelligent robot is installed in the greenhouse to care and help farmers for the melon harvesting.

In Scaffold Mode, the concept of HRI is observed clearly. The human operator and the robot work as a collaborative unified system in which the vehicle autonomously navigates along the structured trees rows whereas the humans on the vehicle concentrate on performing and doing some activities such as 1) thinning, 2) pruning, 3) harvesting, 4) tying trees to wire. Tree trimming tasks were presented with Bergerman et al. [24] and Freitas et al. [25] (see Fig. 8) in which the humans working on the robot in the Scaffold Mode were able to trim trees more than twice as fast as humans using the traditional approach based on ladder.

5. Other Applications

HRI can be found in other applications, which include service, home use, inventory management, and mining. In addition, robots is used for promoting the education for the typical children, whether in the home or in the schools [26, 27]. Space exploration and UAVs are also some applications.

Fig. 9 [28] shows the robot helps the human operator in mining industry. Using the robots increases the arsenal of the tools that help and support the miners to work in more safe and more efficient environment.

Caitlyn Clabaugh and her team [29, 30] developed the socially assistive robot (SAR) tutoring system (Fig. 10) for supporting the educators efforts to teach the number concepts to the preschool children. This system was designed iteratively with the input of the education experts for being developmentally appropriate. The system was investigated and tested in the real-world preschool classroom. The collected data were used for training personalized models of the number concepts learning, leveraging the multimodal data, domain knowledge, and also learning style.

(a) (b)

Fig. 6. The Thorvald robot in the evaluation environment interacting with the pickers: (a) in the open strawberry fields, (b) in the polytunnel environment [22].

Fig. 7. The robot is installed in the greenhouse for helping the farmers for harvesting the melon [23].

Fig. 8. Human operators performing tree trimming while standing in a robot [24, 25].

To achieve highly efficient human-robot cooperation for all applications discussed above, safety system should be found and implemented in the robotic system. In addition, adapted controllers should follow the human collaborator intention and the environment changes. Therefore, this can lead to human-friendly robots.

Fig. 9. Robot helps human in mining industry [28].

Fig. 10. The tutoring system using robot for supporting the educators efforts to teach the number concepts to the preschool children [29, 30].

6. Conclusions

This manuscript has summarized some real applications of HRI. HRI could be found in industrial applications in picking and placing in the production lines, welding processes, assembling parts, and painting. Assistive robotics are one of the highest profile areas of HRI. For people with the physical and

the mental disabilities, robots provide the opportunity for the interaction and the therapy. In addition, HRI could be commonly be widely applied in hospitals and nowadays HRI is very crucial for fighting against the new coronavirus (COVID-19). In agriculture, HRI helps with many tasks such as harvesting, seeding, fertilizing, spraying, hauling, weed detection, and mowing. HRI can also be found in other applications such as service, home use, inventory management, mining, education, and space exploration.

Acknowledgements

The authors would like to thank Dr. Ahmed Elsheikh, Assistant Professor, Mechanical Engineering Department, Faculty of Engineering, South Valley University, Qena, Egypt, for revising and checking the English writing of the manuscript.

References

[1] K. Dautenhahn, Methodology & themes of human-robot interaction: A growing research field, *Int. J. Adv. Robot. Syst.*, Vol. 4, Issue 1, 2007, pp. 103-108.

[2]. M. A. Goodrich, A. C. Schultz, Human-robot interaction: A survey, *Found. Trends® Human-Computer Interact*., Vol. 1, Issue 3, 2007, pp. 203-275.

[3]. W. Knight, Smart Robots Can Now Work Right Next to Auto Workers, *MIT Technology Review*, September 17, 2013.

[4]. J. Kruger, T. K. Lien, A. Verl, Cooperation of human and machines in assembly lines, *CIRP Ann. – Manuf. Technol*., Vol. 58, 2009, pp. 628-646.

[5]. C. Liu, M. Tomizuka, Algorithmic safety measures for intelligent industrial co-robots, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'16)*, 2016, pp. 3095-3102.

[6]. L. Peternel, W. Kim, J. Babic, A. Ajoudani, Towards ergonomic control of human-robot co-manipulation and handover, *IEEE-RAS Int. Conference Humanoid Robot*., October 2017, pp. 55-60.

[7]. C. Thomas, B. Matthias, B. Kuhlenkötter, Human-robot collaboration – New applications in industrial robotics, in *Proceedings of the International Conference in Competitive Manufacturing (COMA'16)*, 2016, pp. 1-7.

[8]. A. Billard, B. Robins, J. Nadel, K. Dautenhahn, Building Robota, a mini-humanoid robot for the rehabilitation of children with autism, *Assist. Technol*., Vol. 19, Issue 1, 2007, pp. 37-49.

[9]. B. Robins, P. Dickerson, P. Stribling, K. Dautenhahn, Robot-mediated joint attention in children with autism: A case study in robot-human interaction, *Interact. Stud*., Vol. 5, Issue 2, 2004, pp. 161-198.

[10]. I. Werry, K. Dautenhahn, B. Ogden, W. Harwin, Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy, in *Proceedings of the International Conference on Cognitive Technology (CT'01)*, 2001, pp. 57-74.

[11]. C. Breazeal, Emotion and social humanoid robots, *Int. J. Hum. Comput. Stud.*, Vol. 59, 2003, pp. 119-155.

[12]. C. Breazeal, Toward sociable robots, *Rob. Auton. Syst*., Vol. 42, 2003, pp. 167-175.

[13]. K. Dautenhahn, Socially intelligent robots : dimensions of human – Robot interaction, *Phil. Trans. R. Soc. B*, Vol. 362, Issue 1480, 2007, pp. 679-704.

[14]. D. Feil-Seifer, M. J. Mataric, Defining Socially Assistive Robotics, in *Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics (ICORR'05)*, July 2005, pp. 465-468.

[15]. T. Kanda, T. Hirano, D. Eaton, H. Ishiguro, Interactive robots as social partners and peer tutors for children: a field trial, *Human-Computer Interact*., Vol. 19, Issue 1, 2004, pp. 61-84.

[16]. P. S. Lum, C. G. Burgar, P. C. Shor, M. Majmundar, M. Van der Loos, Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke, *Arch. Phys. Med. Rehabil*., Vol. 83, 2002, pp. 952-959.

[17]. P. Rani, N. Sarkar, C. A. Smith, J. A. Adams, Anxiety detection for implicit human-machine interaction, in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'03)*, 2003.

[18]. I. Moon, S. Joung, Y. Kum, Safe and reliable intelligent wheelchair robot with human robot interaction, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'02)*, 2002, pp. 3595-3600.

[19]. H. A. Yanco, Development and testing of a robotic wheelchair system for outdoor navigation, in *Proceedings of the Conference of the Rehabilitation Engineering and Assistive Technology Society of North America*, 2001, pp. 1-3.

[20]. R. G. C. Antunes, A robotic system for musculoskeletal rehabilitation of the shoulder biomedical engineering, MS Thesis, *Universidade de Lisboa*, Lisboa, Portugal, 2016.

[21]. COVID-19 test robot as a tireless colleague in the fight against the virus, https://www.kuka.com/en-de/press/news/2020/06/robot-helps-with-coronavirus-tests

[22]. P. Baxter, G. Cielniak, M. Hanheide, P. From, Safe human-robot interaction in agriculture, in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI'18)*, 2018.

[23]. Smart robot installed inside the greenhouse. For the care and help farmers harvest the melon, smart farm on farming 4.0 concept, https://www.shutterstock.com/image-photo/smart-robot-installed-inside-greenhouse-care-765510412

[24]. M. Bergerman et al., Robot farmers: Autonomous orchard vehicles help tree fruit production, *IEEE Robot. Autom. Mag.*, Vol. 22, Issue 1, 2015, pp. 54-63.

[25]. G. Freitas, J. Zhang, B. Hamner, M. Bergerman, G. Kantor, A low-cost, practical localization system for agricultural vehicles, *Lecture Notes in Computer Science*, Vol. 7508, 2012, pp. 365-375.

[26]. M. Cooper, D. Keating, W. Harwin, K. Dautenhahn, Robots in the classroom – Tools for accessible education, in Assistive Technology on the Threshold of the New Millennium (C. Buhler, H. Knops, Eds.), *ISO Press*, Düsseldorf, Germany, 1999, pp. 448-452.

[27]. J. Han, M. Jo, S. Park, S. Kim, The educational use of home robots for children, in *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN'05)*, 2005, pp. 378-383.

[28]. Underground Robots: How Robotics Is Changing the Mining Industry, https://eos.org/features/underground-robots-how-robotics-is-changing-the-mining-industry

[29]. C. Clabaugh, K. Tsiakas, M. Matarić, Predicting preschool mathematics performance of children with a socially assistive robot tutor, in *Proceedings of the Synergies between Learning and Interaction Workshop (IROS'17)*, 2017, pp. 24-28.

[30]. Computational Personalization of Human-Machine Interaction, Machine Learning methods for autonomous adaptation of Human-Machine Interaction to the unique needs and preferences of users over time, http://caitlynclabaugh.info/

(030)

From Hand to Brain and Back: Grip Forces Deliver Insight into the Functional Plasticity of Somatosensory Processes

Birgitta Dresp-Langley
ICube Lab, UMR 7357, Centre National de la Recherche Scientifique CNRS, 67000 Strasbourg, France
E-mail: birgitta.dresp@unistra.fr

Summary: The human somatosensory cortex is intimately linked to other central brain functions such as vision, audition, mechanoreception, and motor planning and control. These links are established through brain learning, and display a considerable functional plasticity. This latter fulfills an important adaptive role and ensures, for example, that humans are able to reliably manipulate and control objects in the physical world under constantly changing conditions in their immediate sensory environment. Variations in human grip force are a direct reflection of this specific kind of functional plasticity. Data from preliminary experiments where wearable wireless sensor technology (sensor gloves) was exploited to measure human grip force variations under varying sensory input conditions (eyes open or shut, soft music or hard music during gripping) are discussed here to show the extent to which grip force sensing permits quantifying somatosensory brain interactions and their functional plasticity. Experiments to take this preliminary work further are suggested. Implications for robotics, in particular the development of end-effector robots for upper limb movement planning and control, are brought forward.

Keywords: Somatosensory brain interactions, Functional plasticity, Human grip force, Sensory input, Visual, Auditory, End-effector robots.

1. Introduction

Neuronal activity and the development of functionally specific neural networks in the brain are modulated by sensory signals. The somatosensory cortical network [1] in the primate brain refers to a neocortical area that responds primarily to tactile stimulation of skin or hair. This cortical area is conceptualized in current state of the art [2, 3] as containing a single map of the receptor periphery, connected to a cortical neural network with modular functional architecture and connectivity binding functionally distinct neuronal subpopulations from other cortical areas into motor circuit modules at several hierarchical levels [1-4]. These functional modules display a hierarchy of interleaved circuits connecting, via inter-neurons in the spinal cord, to visual and auditory sensory areas, and to motor cortex, with feed-back loops and bilateral communication with the supraspinal centers [4, 5]. This enables a 'from local to global' functional organization with plastic connectivity patterns that are correlated with specific behavioral variations such as variations in motor output or grip force. This functional connectivity fulfills an important adaptive role and ensures that humans are able to reliably grasp and manipulate objects in the physical world under constantly changing conditions in their immediate sensory environment. In an exploratory study, a wireless sensor system was exploited to measure human grip force under varying visual and auditory sensory input conditions to probe the extent to which the grip force variations may reveal adaptive changes in somatosensory-visual-auditory interactions.

2. Materials and Methods

A specific wearable sensor system in terms of one glove for each hand with inbuilt Force Sensitive Resistors (FSR), developed in our lab, was used. Hardware and software configurations are described in detail elsewhere here [6, 7]. For the details, see: https://www.mdpi.com/1424-8220/19/20/4575/htm.

The glove was designed to acquire analog voltage signals provided by each FSR every 20 milliseconds (msec) at a 50 Hz rate. In every loop of the Arduino running software, input voltages were merged with time stamps and voltage levels in millivolt (mV). This data package was sent to the computer via Bluetooth, and a specific software wrote the data into separate text files for each individual participant. Data from eight men, between 20 and 30 years old, all of them right-handed, were analyzed. Handedness was confirmed individually using the Edinburgh inventory. Participants were healthy volunteers and gave their informed consent in conformity with the Helsinki Declaration. Each participant was tested individually, standing upright, with both eyes open or blindfolded, facing a table on which two handles, weighing one kilogram each, for power grip exercise were placed in alignment with the forearm motor axis. In the blindfolded condition, they were made to probe for the position of the handles with their two hands before grabbing them. All participants were instructed to "grab the handles with your two hands and gently move them up and down for ten seconds as soon as the music starts." A CD player with constant output intensity played soft yoga music in one test session, and hard rock music in another.

3. Results

Data from the sensor located on the middle phalanx of the middle finger only are analyzed here, given that this finger provides the major contribution to total hand grip force when lifting objects [8]. With one signal every 20 msec, ten seconds per test condition, two visual and two auditory conditions and eight participants, we have a total of 16 000 data for this sensor, and 500 data per individual participant and test condition.

The results (Fig. 1) show clear effects of vision and sound on the middle finger grip forces with blind grip producing stronger forces by comparison with grip eyes open, and hard sound producing stronger grip forces by comparison with soft sounds. The results of 2-Way ANOVA, computed in *Systat* for *Sigmaplot12* on the average middle finger grip forces, are shown in Table 1. The effects of the sound and vision factors are statistically significant, their interaction is not.

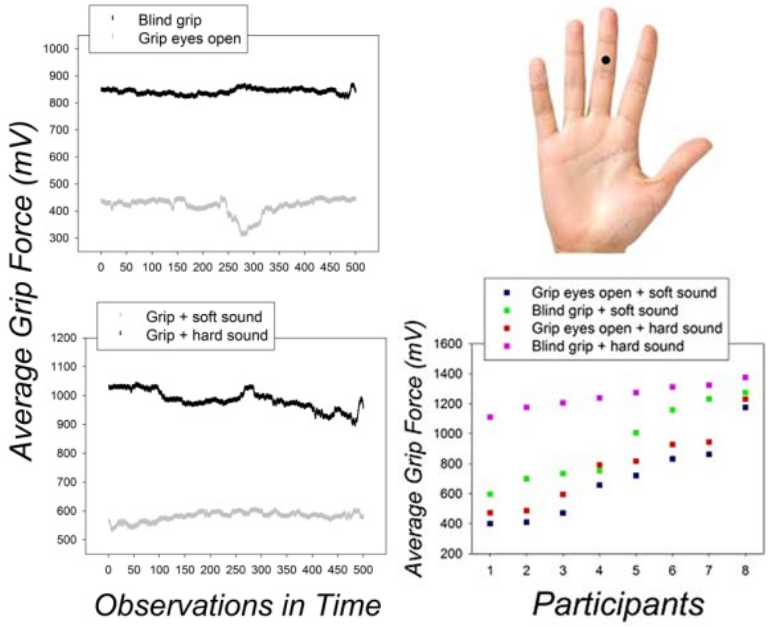


Fig. 1. Average middle finger grip force in the two visual conditions (top left), the two auditory conditions (bottom left), and different individuals' average middle finger grip force in each of the four vision-sound factorial combinations (bottom right).

Table 1. Results of the 2-Way ANOVA with F statistics and their associated probability limits (P).

Two Way Analysis of Variance – Cartesian Design: P8 x V2 x S2 **Dependent Variable:** Average Grip Forces in **Sensor S5**					
Factor	**DF**	**SS**	**MS**	**F**	**P**
Vision	1	1006903.602	1006903.602	18.572	<.001
Sound	1	339897.445	339897.445	6.269	<.018
Vision x Sound	1	104309.710	104309.710	1.924	.176
Residual	28	1518065.682	54216.631		
Total	31	2969176.440	95779.885		

4. Discussion

The middle finger produces most of the grip force necessary for a power grip. Here, it is shown that middle finger grip force adapts spontaneously to changes in the immediate visual and/or sound environment of humans, which translates into a statistically significant, stimulus-driven, change in force output. The spontaneous grip force adaptation is a directly observable consequence, or behavioral correlate, of functional plasticity in the somatosensory brain and connected areas enabling multisensory neural processing where inputs from different sensory modalities are integrated to produce effective sensory-motor activation under changing conditions. This has implications for robot-driven motor rehabilitation through end-effector systems with adaptive multisensory integration simulating the characteristics of human eye-hand coordination with sensory feedback [9-11].

5. Conclusions

Human grip force adaptation is a directly observable consequence of somatosensory brain plasticity. Simpler task designs with a well-controlled parameter space and well-defined boundary conditions can help optimize automated behaviour through multisensory cueing for motor learning and control through human-system-like adaptive feed-back.

References

[1]. S. Wilson, C. Moore, S1 somatotopic brain maps, *Scholarpedia*, Vol.10, Issue 4, 2015, 8574.
[2]. C. Braun, et al., Dynamic organization of the somatosensory cortex induced by motor activity. *Brain*, Vol. 124, Issue 11, 2001, pp. 2259-2267.

[3]. S. Arber. Motor circuits in action: specification, connectivity, and function, *Neuron*, Vol. 74, Issue 6, 2012, pp. 975-89.
[4]. M. Tripodi, S. Arber, Regulation of motor circuit assembly by spatial and temporal mechanisms, *Curr. Opin. Neurobiol.*, Vol. 22, Issue 4, 2012, pp. 615-623.
[5]. T. Weiss, et al., Rapid functional plasticity of the somatosensory cortex after finger amputation, *Experimental Brain Research,* Vol. 134, Issue 2, 2000, pp. 199-203.
[6]. M. de Mathelin, F. Nageotte, P. Zanne, B. Dresp-Langley, Sensors for expert grip force profiling: Towards benchmarking manual control of a robotic device for surgical tool movements, *Sensors*, Vol. 19, Issue 20, 2019, 4575.
[7]. B. Dresp-Langley, F. Nageotte, P. Zanne, M. de Mathelin, Correlating Grip force signals from multiple sensors highlights prehensile control strategies in a complex task-user system, *Bioengineering*, Vol. 7, Issue 4, 2020, 143.
[8]. M. L. Latash, V. M. Zatsiorsky, Multi-finger prehension: Control of a redundant mechanical system, *Adv. Exp. Med. Biol.*, Vol. 629, 2009, pp. 597-618.
[9]. F. Molteni, G. Gasperini, G. Cannaviello, E. Guanziroli, Exoskeleton and end-effector robots for upper and lower limbs rehabilitation: narrative review, *PM&R*, Vol. 10, Issue 9, Suppl. 2, 2018, pp. S174-S188.
[10]. L. P. Wijesinghe, J. Triesch, B. E. Shi, Robot end effector tracking using predictive multisensory integration, *Front. Neurorobot.*, Vol. 12, 2018, 66.
[11]. T. Schürmann, B. J. Mohler, J. Peters, P. Beckerle, How cognitive models of human body experience might push robotics, *Front. Neurorobot.*, Vol. 13, 2019, 14.

(031)

Formation Control of Five Vehicles for Load Transportation Under Randomly-switching Communication Topology

A. Tahri [1] and L. Guenfaf [2]
[1] LRPE Laboratory, University of Science and Technology Houari Boumediene (USTHB),
BP 32 Alia, Algiers 16111, Algeria
[2] LSEI Laboratory, University of Science and Technology Houari Boumediene (USTHB),
BP 32 Alia, Algiers 16111, Algeria
E-mails: a.tahri@usthb.dz, lakhdar.guenfaf@usthb.dz

Summary: the objective of this work is to validate a formation control algorithm ensuring that a group of 5 vehicles achieve in the shortest possible time a prespecified shape formation and maintain it while moving towards their final goals. First-order integrator models were used to develop the control law. The effect of switching communication topologies is considered in simulations by randomly switching between different communication topologies at different instants during formation stabilization and formation tracking. Matlab simulations showed the effectiveness of the adopted algorithm.

Keywords: Cooperative control, Formation control, Multi-vehicle system.

1. Introduction

Formation control is a very active topic in the cooperative control field, usually, two classes of the problem are considered: formation stabilization and formation tracking. Formation stabilization consists of moving a group of vehicles to form a prescribed shape. In formation tracking the whole team is moving to follow a desired trajectory while maintaining the formation shape.

In this paper, we present a validating simulation of a formation tracking control algorithm for five vehicles considering first-order integrator dynamics.

1.1. Literature Overview

Formation control is studied for diverse types of vehicles. In [2-4] different communication graph topologies were tested. The virtual leader technique is combined with a behavioral approach to design a control structure for wheeled mobile robots. In [5] a behavioral-based approach was developed for a box-pushing task. Moreover, the coverage problem is solved with graph-based methods for planetary exploration tasks in [1]. Formation control is studied for UAVs as well. Systems such as multi-spacecraft have also been considered in various configurations. Deepspace spacecraft formation is studied in other works where a ring topology communication graph is considered. The authors proposed a distributed control structure for formation stabilization, the structure is simulated for simplified dynamical models. Some works discussed the communication constraints effects on formation stability.

1.2. Graph Theory Preliminaries

For modeling communication among the vehicles, we use undirected graphs where the vertices represent the vehicles and the communication links between every two vehicles are represented by the graph edges.

A (simple) undirected graph is a pair (ν, ξ) that consists of the vertices set ν and edges set $\xi \subseteq \{(i,j) \in (\upsilon \times \upsilon) : i \neq j\}$ with: $(i,j) \in \xi \Rightarrow (j,i) \in \xi$.

The adjacency matrix of an undirected graph $A = [a_{ij}]$ is a symmetric hollow matrix defined as following: $(i,j) \in \zeta \Leftrightarrow a_{ij} \neq 0$ and $a_{ij} = 0$ otherwise. The set of neighbors of a node i is defined by $N_i = \{j \in \upsilon : a_{ij} \neq 0\}$.

2. Problem Statement

In this section, we present the mathematical formulation of the formation tracking problem to set the scene for the control law development section.

Let $p_i = \{x_i, y_i\}$ represent the coordinates vector for the vehicle i on R^2, with $i \in \upsilon$, the dynamics of a vehicle i are defined by:

$$\dot{p}_i = u_i, \; i = 1,...,n \tag{1}$$

The problem of (shape-based) formation control is stated as follows:

Let p_i, $(i \in \upsilon)$ the coordinates vector of n vehicles modeled by (1), define the control u_i ensuring:

1. $\|p_i(t) - p_j(t)\| \to d_{ij}$ $t \to \infty$, for all $(i, j) \in \xi$

   with d_{ij} is defined upon to the prespecified shape.
2. $\|p_i(t) - p_{g_i}(t)\| \to 0$ as $t \to \infty$ for all $i \in \upsilon$, with p_{g_i} is the final desired position for vehicle i.

3. Control Law Development

To solve the formation control problem defined above, we propose a control law combining two actions: a variant of local information-based consensus law (see [4] for relative state deviations) and a proportional action (ensuring $p_i \to p_{g_i}$)

$$u_i = -coeff_1 * \sum_{i \in \upsilon} a_{ij}(t) * \begin{bmatrix} (p_i(t) - p_j(t) - Coeff_2 * \\ * \sum_{i \in \upsilon} (p_{g_i}(t) - p_{g_j}(t)) \end{bmatrix} + \\ + Coeff_3 * (p_i(t) - p_{g_i}(t)) \quad (2)$$

4. Simulation Results

In this section, we present Matlab simulations of formation tracking achieved by five vehicles modeled with first-order dynamics.

The communication graph topology is supposed to be switching randomly between one of the topologies plotted in Fig. 1.

In Fig. 2 the vehicles' trajectories are presented for different parts of the formation achieving time intervals.

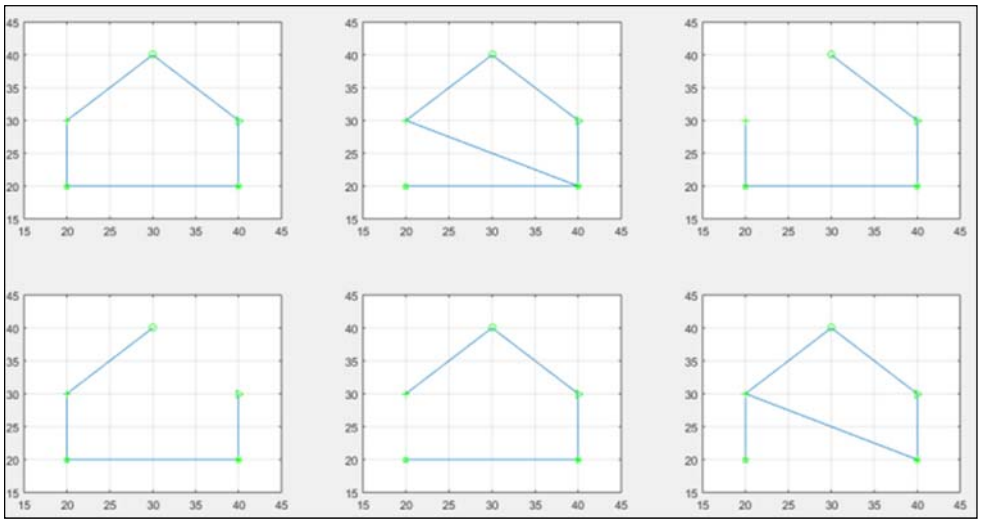


Fig. 1. Different communication graph topologies.

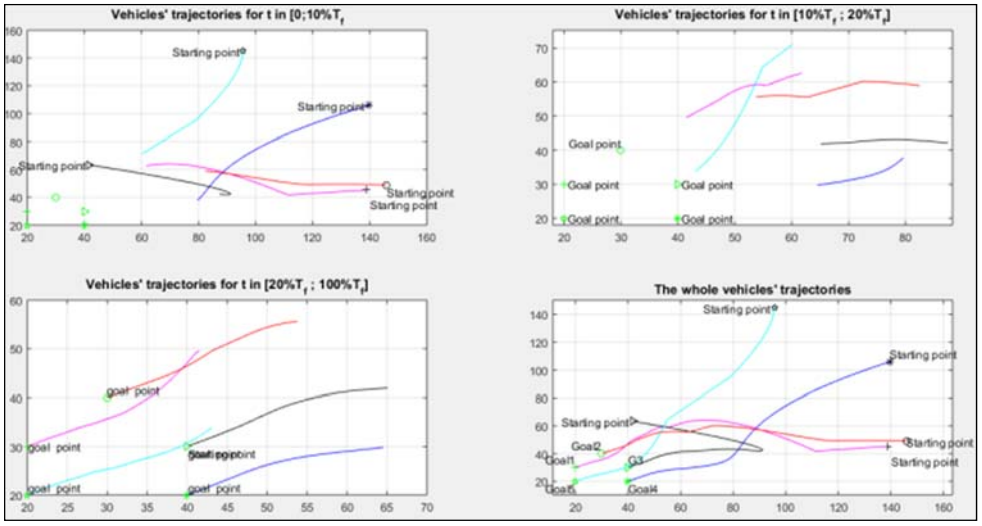


Fig. 2. Trajectories of vehicles plotted for different time intervals, on the top left corner, the trajectories are plotted from starting instant until 10 % of the arrival time, the top left plot shows the trajectories for $t \subset [10\,\%T_f : 20\,\%T_f]$ since when the formation shape is achieved, the last part of the trajectories (for $t \subset [20\,\%T_f : 100\,\%T_f]$ is plotted in the bottom left window, the whole trajectories are plotted in the bottom right window.

5. Conclusions

In this extended abstract, we presented a validation of the formation tracking control algorithm, first-order integrator models were used to develop the control law, Matlab simulation showed the effectiveness of the adopted algorithm.

References

[1]. N. Gyagenda, A. K. Nasir, H. Roth, et al., Coverage path planning for large-scale aerial mapping, in *Proceedings of the Annual Conference Towards Autonomous Robotic Systems (TAROS'19)*, 2019, pp. 251-262.

[2]. Z. Li, D. Zhisheng, Cooperative Control of Multi-Agent Systems: A Consensus Region Approach, *CRC Press*, 2017.

[3]. Z. Peng, D. Wang, H. Zhang, G. Sun, H. Wang, Distributed model reference adaptive control for cooperative tracking of uncertain dynamical multi-agent systems, *IET Control Theory & Applications*, Vol. 7, Issue 8, 2013, pp. 1079-1087.

[4]. W. Ren, R. W. Beard, Distributed Consensus in Multi-Vehicle Cooperative Control (E. D. Sontag, M. Thoma, Eds.), *Springer*, London, 2008.

[5]. C. R. Kube, H. Zhang, The use of perceptual cues in multi-robot box-pushing, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'96)*, Minneapolis, MN, *USA, USA*, 22-28 April 1996, pp. 2085-2090.

(034)

Intelligent Measurements as a Bridge between Measurement Theory and Artificial Intelligence: Bayesian Measurement Neural Networks (BMN) Based on the Methodology of the Regularizing Bayesian Approach

S. Prokopchina
Financial University under the Government of the Russian Federation, Leningradsky pr., 49, Moscow, Russia
Tel.: + 79150070489
E-mail: svprokopchina@mail.ru

Summary: The paper considers the approach and methodological principles of creating a new type of neural networks, called Bayesian measurement networks. The concept and formalization of a new type of Bayesian neurons implementing Bayesian convolution based on the regularizing Bayesian approach is given. Three types of Bayesian neurons that implement convolution of values of quantitative and qualitative features are considered an architectural scheme and metrological justification of BIN solutions are proposed. A platform for rapid development of applied BMN is proposed. Examples of solutions to applied problems based on BMN are given.

Keywords: Neural networks, Bayesian convolution, Regularizing Bayesian approach.

1. Introduction

Neural network technologies are increasingly used in solving complex practical problems. Moreover, hybridization, that is, the integration of neural network technologies with other information technologies, such as technologies of systems based on fuzzy sets, is characteristic of the current stage of their development. This trend is due to the desire to expand the range of tasks solved by neural networks from image processing and text recognition to analytical data processing [1, 2].

The purpose of constructing such systems is to solve the problem of determining the output values of the objective function for certain sets of input data without computational processing. These transformations are accompanied by significant reductions in the dimension of the input arrays of information. However, the traditionally noted disadvantages of neural networks include the impossibility of semantic interpretation of the resulting solutions, managing the network operation, and obtaining intermediate solutions. In addition, the quality of input information and its metrological specifics are not taken into account, and therefore the quality of output solutions is not determined.

These shortcomings can be successfully overcome when creating a neural network structure based on the methodology and technologies of the regularizing Bayesian approach.

It is worth noting that the RBP methodology is developed on the basis of the principles of the theory of pattern recognition, and therefore can be reasonably used to build neural networks.

2. Methodological Foundations and Principles of Creating Bayesian Measurement Neural Networks

The regularizing Bayesian approach (RBP) [3-5] has become one of the successful examples of creating platform solutions for modeling problems and creating developing information technologies, as well as monitoring, auditing and managing them under uncertainty.

The modern concept of a complex object under a deferral on the basis of models with dynamic constraints (MDO) involves the construction of specific models in the process control system of a complex object or behavior object in difficult-predictable situation and allows the possibility of changing how the properties of the object forming the MDO and their relationships and the systems of representation of properties and relations. The results of the solutions are presented as h_{kt} benchmarks of a special type of scales, called conjugate scales with dynamic constraints (SDC) with changing properties according to the incoming information

Let's denote the carrier of such a scale as H_{KT}. Then we can formalize it as:

$$H_{KT} = \{h_{kt}\}; (h_{kt} \subset H_{KT}; k = 1, K; t = 1, T), \quad (1)$$

where K is the number of scale reference points, t is the measurement time, and T is the measurement time period. The view of the SDC is shown in Fig. 1. It is a two-scale structure for placing numerical (on the upper scale) and linguistic (on the lower scale) information. Such scales are called conjugate because they allow

you to implement the convolution of numerical and linguistic information.

Such scales are constructed for each measured property of a complex object and the influencing factors of the external environment. Together, these interconnected scales are a hyper scale (hypercube), which implements a model of the object and its functioning environment in the form of an MDO.

3. A New Type of Neurons Based on Probabilistic Bayesian Convolution

When creating neural networks based on RBP, it can be noted that in fact such a hypercube implements the structure of a convolutional deep learning neural network of a hierarchical type. The convolution is based on the modified Bayesian formula (5).

The first block of the network implements configuration (network training), which determines the types of probability density for numerical data and the types of membership functions for linguistic variables, as well as the MDO, the hierarchy of which determines the structure and number of convolutional layers of the network.

The second block performs the function of fuzzification of variables and metrological justification of decisions based on the following equation:

$$\{h_{kt}|\{MX_{kt}\} = \{\text{argmin } C[\varphi_{jt}(x_{it}|y_{it})]\}, \quad (2)$$

where $\{h_{kt}|\{MX_{kt}\}$ is the fuzzified value of the input variable with complexes of metrological characteristics of accuracy, reliability, and reliability for each component of the fuzzy set; C is the optimization criterion in the form of the average Bayesian risk of the solution; φ_{jt} is the Bayesian fuzzification function; $x_{it}|y_{it}$ is a set of data obtained under y_{it} conditions, including sets of a priori information, constraints, and metrological requirements.

The chain of constant metrological support for each stage of the implementation of the network algorithm allows us to reasonably define this type of neural network as a ***measuring neural network***. And since a modified Bayesian rule (5) is used to implement convolution, such a network can be called a ***Bayesian measurement neural network*** (BMN) of a hierarchical type with deep learning.

With this architecture, convolutional layers can be represented as feature convolution layers, which allows you to interpret the decisions of each layer in relation to their semantic content.

When convolving two or many features, each neuron of such a network implements on its convolutional layer an optimizing Bayesian function of the minimum average risk of a solution with the form:

$$\{h_{kt}|\{MX_{kt}\} = \{\text{argmin}C[\varphi_{jt}(x_{it}|y_{it}; x_{i+1t}|y_{i+1t})]\}, \quad (3)$$

where $h_{kt} \subset H_{KT}$; k = 1,K; t = 1,T; φ_{jt} – Bayesian convolutional function; $\varphi_{jt} \subset \Phi_{JT}$; j = 1, J; $x_{it} \subset X_{IJT}$; (i = 1, I) - sets of values of numerical and linguistic variables reflecting the properties of features; Y_{it}, Y_{IJT}; (i = 1, I) – conditions for obtaining data and implementing convolution algorithms; Y_{IJT} = {A_{IJT} - set of a priori information; O_{IJT} - set of constraints; MX – Multiple Metrological requirements}

The Bayesian convolution formula for the convolution of a priori and incoming information has the form:

$$P(h_k \mid x_i \mid Y_i) = \frac{P\ (h_k \mid Y_{i-1}) l(x_i \mid h_k \mid Y_i)}{\sum_{j=1}^{K} P\ (h_j \mid Y_{i-1}) l(x_i \mid h_j \mid Y_i)} \quad (4)$$

The Bayesian convolution formula for the convolution of two features can be represented as follows:

$$P(h_{kt}^{F_t} \mid G_t^{(OE)}) = \\ = \frac{P^a(h_{kt-1}^{(F_t)} \mid G_{t-1}^{(OE)}) \overset{L}{\circ} P(G_t^{(OE)} \mid h_{kt}^{(F_t)})}{\sum_{j=1}^{J} P^a(h_{jt-1}^{(F_t)} \mid G_{t-1}^{(OE)}) \overset{L}{\circ} P(G_t^{(OE)} \mid h_{jt}^{(F_t)})}, \quad (5)$$

where P(.) denote the a posteriori probability values of the reference points h numeric or linguistic scales of the type SDC for factor F in time t; $P(G^{(OE))}$ – posterior probability values of factors in the external environment.

In the process of such convolution, the probabilities of the benchmarks of the feature scales are transformed, information about which enters the system in the form of input data streams and knowledge about the features.

Let's call this type of neuron ***a Bayesian neuron.***

Following these aspects and taking in consideration metrological support of receiving decisions, the results of BMN are named as neuronet measurements.

Let's take a closer look at the implementation of Bayesian convolution to explain the above definitions and the above reasoning.

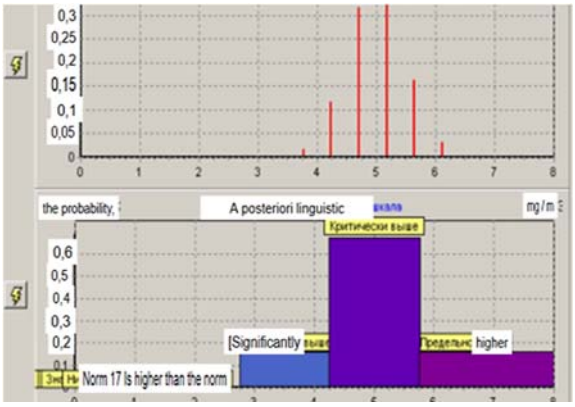


Fig. 1. The conjugate scale of the SHDO type used to implement Bayesian convolution.

At the heart of the synthesis of information technologies of RBP, the principle of unity of measurements is applied, which allows to coordinate the inputs and outputs of individual scales and transform them in accordance with the functional content of information technology and compliance with metrological requirements for information system solutions. For this purpose, in parallel with the computational process, the process of metrological support of each solution is implemented in the form of indicators of accuracy, reliability, reliability, entropy and risk. These indicators are combined into complexes of metrological characteristics.

The solution obtained on the basis of Bayesian convolutions is a series of alternative feature estimates with corresponding complexes of metrological characteristics and is a regularized Bayesian estimate (RBE), the properties of which are considered in [3].

In the BIN, according to the neuron formula, three types of convolutions are implemented:

1. Convolution of a priori and a posteriori values of numerical or linguistic variables (Bayesian neuron type 1);
2. Convolution of the values of numerical and linguistic variables (Bayesian neuron type 2);
3. Convolution of the integrated values of various features (Bayesian neuron type 3).

4. Software Implementation of the BIN

The presence of a system of metrological support allows to create a neural network with the ability to manage the quality of the resulting solutions.

With multiple implementation of convolution, there is a sharp decrease in the dimension of the feature space, which allows processing a significant number of data streams at high speed. On these principles we built the platform 'Infoanalyst' which is essentially a BMN. Unlike neural networks in application systems built on its basis, all convolutional layers are easily viewed in the user interface and interpreted using cognitive graphics. For each feature, as solutions, you can get comprehensive information, dynamic characteristics, trends, assessments of situations and recommendations for their adjustment in accordance with the specified norms and rules. An example of BIN solutions for heat power engineering is illustrated in Figs. 2 and 3.

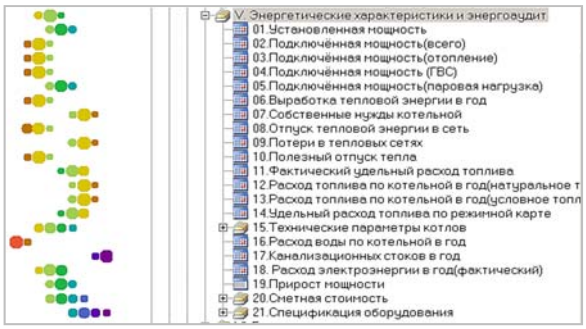


Fig. 2. Convolution of indicators of heat-generating systems with cognitive interpretation of solutions.

Other numerous examples of applied solutions can be found in [3-5].

5. Conclusion

The proposed neural network model has unique properties, including interpretability and cognitive visualization of the resulting final and intermediate solutions, manageability of the quality of solutions, the ability to explain solutions, as well as the development of the network structure and functioning in conditions of significant information uncertainty. An important feature of the network is the ability to process both data and diverse knowledge, which is necessary in conditions of uncertainty. In addition, an important property of the BIN is the ability to convolve data and knowledge about the measured property, which allows you to obtain additional information. Therefore, for this type of neural network, there are no restrictions and difficulties in developing a data set and training the network. Given all of the above unique properties of the BIN, the network can be recommended for use in Industrial 4.0 tasks.

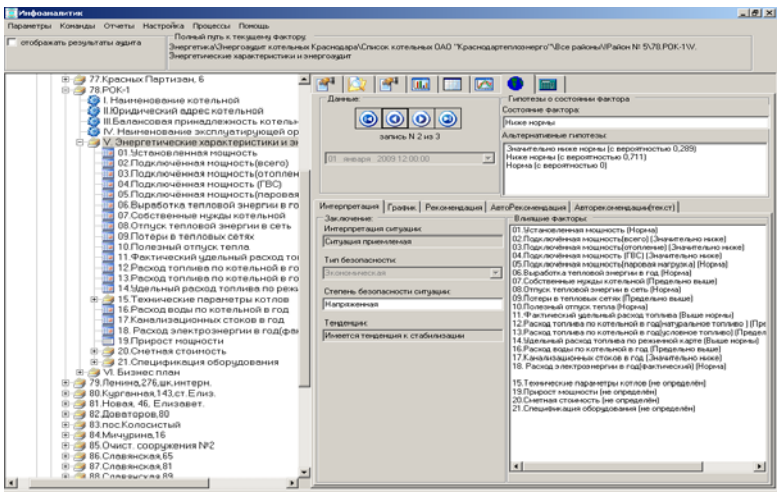


Fig. 3. Convolution of signs of the state of the energy-generating system with explanations.

Reference

[1]. N. G. Yarushkina, Fundamentals of the Theory of Fuzzy and Hybrid Systems, *Finance and Statistics*, 2004.

[2]. M. N. Koroleva, Monitoring of a complex technical object based on cognitive measurement, PhD Thesis 2018.

[3]. S. V. Prokopchina, Soft measurements and management of complex systems based on RBP, in Soft Computing and Measurements, *Publishing House "Scientific Library"*, Vol. 1, 2017, pp. 3-89.

[4]. S. V. Prokopchina, Bayesian intelligent technologies based on RBP, Management in Conditions of Uncertainty (S. V. Prokopchina, M. Yu. Shestopalov, Eds.), *LETI Publishing House*, St. Petersburg, 2014, pp. 3-78.

[5]. S. V. Prokopchina, Cognitive measurement networks on the base of regularizating Bayesian approach, Soft Measurement and Computing, *Publishing House "Scientific Library"*, Vol. 9, 2018, pp. 7-25.

(036)

Providing Measurement Trustworthiness is the Key to Industry 4.0 Realisation

K. Sapozhnikova, A. Pronin and R. Taymanov
D. I. Mendeleyev Institute for Metrology, 19 Moskovsky pr., 190005 Saint Petersburg, Russia
Tel.: + 79213066126, fax: + 78127130114
E-mail: k.v.s@vniim.ru

Summary: The technologies that determine the pace of society development are driving the use of measuring instruments (MIs) and multichannel measuring systems (MMSs) in these technologies. The need to reduce costs forces combining the technology development with the search for effective methods and instruments for providing measurement trustworthiness. The analysis of development trends of this process gave grounds to propose a procedure for checking the metrological health of MIs and MMSs in the course of their operation, which was called "metrological self-check". It is expedient to supplement this procedure with virtual tests. As a result, calibration/verification intervals of MIs and MMSs can be extended. With artificial intelligence use, it becomes possible for traditional measurements to fully automate the procedure for providing measurement trustworthiness. The tendency of increasing requirements for assurance of soft measurement trustworthiness is shown.

Keywords: Measurement trustworthiness, Metrological self-check, Virtual tests, Digital twins, Artificial intelligence, Soft measurement.

1. Introduction

The conception of the need to ensure the trustworthiness of measurements has emerged at the early stage of the legal metrology development. Measurements are a meaningful form of communication; it can be communication of people with each other or, in a figurative sense, with the process they are controlling. The discrepancy between a measured value expected by its "recipient" and a real quantity value measured, can lead to undesirable situations, sometimes with tragic consequences.

Any measuring instrument (MI) is characterised by a certain calibration curve that expresses "the relation between indication and corresponding measured quantity value" [1]. In case of a measuring transducer or material measure, the value specified in documentation is used instead of indication.

Therefore, a fundamental metrology requirement aimed at providing the trustworthiness is as follows: for any MI, the calibration curve should remain within a permissible uncertainty limits under specified external conditions throughout the MI lifetime. As a consequence, to check the fulfilment of this requirement, i.e., to check MI metrological health, is necessary.

Since over time, characteristics of any MI are changing due to originated defects, systematic worsening of material properties, or, e.g., weakening component fasteners, the concepts have originated related to periodic calibration / verification, and intervals between such procedures.

These procedures involve the cost of their implementation. In the course of human civilization development, methods of measurements and MI technologies have been improved, the number and accuracy of MIs being increased. Accordingly, the total costs have increased, and it has been desirable to reduce them. As a result, the evolution process of trustworthiness provision methods has become an organic part of technology development.

The pace of technology development at the second part of the 20th century and beginning of the 21st century is many times higher than that inherent in the earlier period. The analysis of correlations between technology development stages, on the one hand, and features of the methods for measurement trustworthiness provision, which were developed during these stages, on the other hand, can help to reveal the regularities of the considered process. Looking under this angle, to foresee new tasks in this field, which specialists in metrology will face in the near and more distant future, becomes possible.

2. Provision of Measurement Trustworthiness as Organic Part of Technology Development in the 20th Century

In this context, the most meaningful technological solution that has been applied since the beginning of the 20th century, is conveyor production with the elements of control automation.

Significantly, any forced conveyor shutdown and the resulting downtime in its operation lead to the costs due to the output reduction, while the occurrence of a metrological failure in used MIs results in supplement costs related to production spoilage.

As a consequence, the requirements for checking the metrological health of MIs applied in technological processes, including those embedded in equipment, have arisen. Ideally, the procedure should be carried out after each measurement or a small measurement series being fulfilled by the instrument under checking,

but without stopping the equipment. A delay between the measurement and checking procedure may take place, the delay time being limited due to possible economic losses permitted for this delay time.

To carry out such a periodic checking is possible based on indirect features, in particular, production features, e.g., by measurements of the finished part sizes which should correspond to the specified values [2].

In the years that followed, and up to the present day, these technological solutions have been improved and used more and more widely in automatic machinery, power units, transport, including aviation and spacecraft, in medical technology, etc. The metrological failure consequences value has been growing up. At present, such failures can lead to human losses. Accordingly, the requirements for checking have been tightened. In technical systems of high importance, calibration / verification intervals should be reduced to significantly lower values than the calculated ones.

At the same time, the human role in automatic control systems was concentrated on the equipment operation failures, including those caused by appearance of untrustworthy measurement information and on the prevention of situations which could result in accidents.

The search for new solutions has led to the understanding that the prototype of automatic control systems (ACS) first emerged in living organisms and that evolutionary experience could be useful in improving technologies. Cybernetics has emerged, drawing on the analogy between the biological evolution and development of technical systems. Then, computer technologies have become a part of production process, which provided tools for automatic control system improvement.

At the end of the 20th century, the rapid growth in the number of MIs and multichannel measuring systems (MMSs) included in ACS was noticeable. It was accompanied by the transmission of measurement information over significant distances in order to carry out control operations and by the use of computer technologies including cloud-based ones for processing measurement information.

These tendencies have stimulated the search for methods that allow minimizing costs for provision of measurement trustworthiness by transition to automatic checking the metrological health (metrological self-check / self-validation) of MIs and MMSs [2-7]. Eventually, this way will lead to the reduction in the number of metrology specialists performing periodic metrological maintenance.

Methods of metrological self-check are based on the application of redundant information sources that are available or artificially introduced in MIs and MMSs. They provide the ground for essential extension of calibration / verification intervals, and in the event of metrological failure, enable taking measures to eliminate it quickly with the minimum costs.

The end of the 20th century has brought strengthening the links between professionals living in various countries. People were increasingly treated or studied abroad. The tendency has been marked to integrate processes across sciences in the new technology development including those related to the re-engineering of clinical medicine. The need for metrological traceability has emerged not only in engineering and physics, but also in chemistry, laboratory medicine, biology, biochemistry, molecular biology, food science, forensic science, etc. [1].

Since 1984, three editions of the International Vocabulary of Metrology have been published which reflected required changes. The fourth edition is currently under development.

3. Expected Changes in Technology and Associated Metrology Solutions at Industry 4.0 Stage

Understanding the relevance of extending the boundaries of the measurable was clearly shown at the last IMEKO TC 1 – TC 7 – TC 13 – TC 18 Symposium held in Saint Petersburg in 2019. About one third of papers presented there related to the measurements of quantities that before were considered as qualitative properties.

The same trend is confirmed by the analysis of the Scopus data. The proportions of publications dealing with measurements in special fields and those related to measurements in all fields were calculated. In 2001-2019, this proportion increased by 1.7 times in what concerns measurements in education, twice in the case of sociology, by 2.3 for zoology, and by 5.5 for psychology.

The range of mass-produced MIs, especially for chemical manufacturing processes and medicine, is noticeably expanded. Among the latter, MMSs automatically forming a diagnosis based on many parameters, are of particular interest, e.g., [8]. Accordingly, a variety of methods and means for ensuring measurement trustworthiness increases.

At Industry 4.0 stage, the vast majority of computerised MIs and MMSs will form the basis for the ACS of cyber-physical complexes that usually include "smart manufacturing", "smart transport", "smart houses", "smart cities", "smart healthcare", etc. The exploitation life of such complexes before being modernised is highly likely to be estimated not less than 10 years. With the present-day instrumentation technologies, there exists a significant probability that the metrological characteristics of the MIs and MMSs used in cyber physical complexes will lie outside the permissible limits before they are modernised.

Among others, cyber-physical complexes include centralised supply systems for such resources as gas, electricity, water, and petroleum products. The need for controlling resource flows and measuring their parameters has led to the emergence of appropriate ACSs with MMSs distributed over large areas. In

many cases, measurement results obtained in MMS channels are directly or indirectly interlinked. Then, there exists a possibility to organise metrological self-checking based on the available functional redundancy.

In such systems, as resources are being transmitted to consumers, their amount, density (in case of precipitations), chemical composition (in case of transmission breach), etc. can change. The metrological self-check enables quickly determining problems originated as well as their possible reasons with the minimum costs. After that, the methods for troubleshooting can be found.

In particular, to estimate the required resource amount in various districts and maintain ACSs in a serviceable condition, the data on resource consumption received by each customer which is confirmed by device indications, should come to control centers. If the data from customers do not correspond to the center data on the resource amount that has come to this district, the measures should be taken on revealing the places where failures occurred. In many cases, possible reasons for such a situation concern MI metrological failures.

In order to organize the efficient metrological self-check, MMS should be multilevel, each channel from the center at any level being divided into next level channel groups transmitting the resource to a part of users. The metrological self-check can be performed with the help of subsystems, that link the input MIs of channels related to a previous level with the input MIs of subsequent level channels. This subsystem should carry out checking the correlation between the amount of the transmitted resource and the data on summarized resource received by a group of consumers.

In case of discrepancy, the results of such a check will enable automatic localising an unhealthy MI or will give grounds to draw the conclusion that, for example, the reason of the revealed discrepancy is the initiation of the leakage in the channel between the MIs of previous and subsequent levels.

A failure signal that indicates the necessity for a specialist to come and take corrective maintenance should be sent to the center.

However, in such MMSs, the probability exists that this information will not come to the center in time due to a too long cycle of message collection from all MMS channels and to malfunction of the message collection sequence.

In order to exclude this risk and economic losses related to it, when being sent to the center, each message should receive a priority rank [9], information on more significant failure having a higher one. In this case, the measures for eliminating an occurred failure will be taken promptly with the minimum losses.

Such ACSs with various automation level exist in a number of cities and districts, they continue to be improved, but no publication could be found anywhere about the implementation of the developed metrological self-check in full.

Further steps in the development of methods for trustworthiness provision cross the boundaries of the analogies between technical solutions and those found in the course of biological evolution [5]. The search moves into the realm of analogies with methods developed along the path of technical civilization.

A characteristic example is the choice of solutions based on individual experience. It was required in the implementation of metrological diagnostic self-checking. It was exactly this method that has brought to the use of new metrology concepts, in particular, "critical uncertainty component" error "virtual tests" [5, 6, 10-12].

Critical uncertainty component is a predominant component or that inclined to fast growth. (In some cases, it can be the sum of several components.) To determine it, on the first stage, to analyse the operation experience of an analog (a MI with similar design) or to produce a MI or MMS channel prototype, is necessary. Then, this prototype or its analog should be tested for the impact of influence quantities that are considered significant. On the second stage, to analyse various uncertainty components and single out those that correspond to the definition given above, is necessary. If a part of them can be decreased significantly by changing some element design of the prototype, it should be done. Then, the critical uncertainty component should be singled out. The method of its evaluation should be found; it is just this component that should be applied for the metrological self-check.

The concept of virtual tests has been widely spread in recent years, but to date, it has not been standardised. Virtual tests can be defined as tests carried out with the help of a digital twin [13, 14], i.e., a tool that simulates the state and behaviour of the environment or system affecting the test object.

Below, a version of virtual tests is considered. They are aimed at estimating the metrological reliability related to the part of a MI or MMS channel. This part consists of components that are located outside a sensor with digital output. All transducers included into a channel that transmit measurement information to a data processing unit and the processing unit itself are under test. The latter generates signals for the ACS operator based on the results of metrological self-check.

Significantly, the digital twin here is implemented taking into consideration the sensor characteristics. However, in practice, the used data can be obtained from analogous sensors that have some differences from those planned to be applied in the future. To pay attention to this methodical feature of virtual tests is necessary.

The resulting estimate of metrological reliability should correspond to the value sufficient to realise the required calibration / verification interval of the MI or MMS. Such tests are performed with disconnected sensors. The secondary transducer inputs are connected with devices that contain the records of the most typical fragments representing measured changes and their combinations obtained in the course of a close analog operation. Sometimes, these records are supplemented with dangerous combination fragments

designed as a result of the mathematical simulation, which take into account rare and/or unlikely situations and impacts. Together, the records being applied constitute a digital twin of the object affecting the MI or MMS. The metrological reliability of the sensors is estimated separately.

For specific cases, the interpretation of the conceptions "the most typical fragments" and "dangerous fragment combinations" depends on a subjective decision. This circumstance gives reasons to consider "virtual tests" as a form of "soft" measurements of metrological reliability. Taking into account [15, 16], here, by "soft" the authors mean measurements made under conditions of considerable uncertainty related to a measurement object model.

Since the subjective factors can influence the efficiency of self-check, while devices with the metrological self-check are produced in a number of countries, to issue an international standard is actual. It should include unified requirements for MIs and MMSs based on the accumulated experience of their development and operation as well as national standards and guides issued before.

4. Artificial Intelligence and New Approaches to the Provision of Measurement Trustworthiness

Industry 4.0 has just begun, but the pace of technological development is such that we can already foresee the contours of its completion and start of the next stage.

A key feature of artificial intelligence is adaptability, i.e., the ability to adjust to changing conditions.

Mass production will become customized. The role of artificial intelligence will increase significantly, while human involvement in production processes will decrease and, in some cases, be reduced to nothing. Cyber-physical cognitive systems, including robots, will be widely spread. They will often use procedures that are characteristic of intellectual activities, such as image recognition, analysis of trends and their consequences, prediction, decision-making, etc.

At present, many types of sensors and other MIs are produced which measurement range can be adjusted. However, there have emerged the first sensors with an embedded neural network being capable of searching and selecting optimal parameters that correspond to the input signals and influence quantities.

If they are provided with the metrological self-check function, they can be called intelligent, because according to GOST [17, 18], an intelligent sensor / intelligent MMS is an adaptive sensor / adaptive measuring system, with the function of metrological self-check. In considering this definition, to take into account is necessary that the intelligence can be of different level.

The intelligent sensor, as a rule, should have a digital output and provide transmission of information on a metrological failure occurred.

To increase calibration / verification interval without worsening measurement trustworthiness, self-check methods should be supplemented with automatic metrological correction methods (i.e., methods of compensation for uncertainty changes).

The combination of metrological self-check and self-correction functions brings the implementation of MIs and MMSs realising them closer to sensor systems of mammals (surely, including humans) with advanced intelligence. In fact, this makes the intelligence level of the former ones higher. (The behaviour of people who experience visual or hearing impairment, but do not have a possibility to use glasses or hearing aid, is characteristic. They take possible measures to compensate for the originated defect by increasing the attention to information from other sensory systems that are healthier. Besides, they behave more carefully.)

At the end of Industry 4.0 stage, intelligent sensor with necessary calculation resources will enable revealing metrological failure and also^

- Automatic correcting uncertainty that appeared as a result of influence quantity impact and/or material ageing;
- Self-recovering if a single sudden malfunction has occurred in the sensor;
- Self-learning.

Self-recovery means an automatic procedure that weakens the metrological failure consequences, i.e., it is a procedure for fault tolerance provision, while fault tolerance is the ability to keep metrological characteristics within permissible limits in case of a single defect.

Self-learning implies the ability to automatically optimise parameters and operation algorithms. One of the important purposes of self-learning is developing the reaction to unexpected events.

In a near future, intelligent MMS will include:

- Intelligent sensors;
- Several sensors that measure the same quantity and are located at a distance, one of them being more accurate and reliable than the others;
- Sensors that measure various quantities the relation between these quantities being known with a required accuracy.

Accordingly, besides revealing metrological failure in one or several measuring channels, intelligent MMS will be able (like in case of intelligent sensors) to carry out automatic correction of uncertainty, self-recovery, and self-learning.

If an intelligent sensor or intelligent MMS is provided with the self-learning capability, a corresponding built-in neural network should be able to be checked, i.e., should be able to reproduce algorithms being realized, including those related to occurrence of unexpected situations. Ideally, the neural network will automatically check whether the

solutions made as a result of self-learning correspond to specified limitations.

5. Conclusions

The analysis of tendencies in development of methods and means for providing measurement trustworthiness demonstrates that for traditional types of measurements, organization of the metrological self-check based on the available or artificially introduced redundancy is the most perspective way. In new MIs and MMSs, the metrological self-check should preferably be supplemented with the self-correction of growing uncertainty.

Requirements for terminology, methods and the level of metrological self-check and self-correction should be reflected in an international standard. The conformity between the level implemented in a MI or MMS and the requirements of the standard should be substantiated in their documentation.

Virtual tests have shown to be effective, but there exists a need to systematize the experience gained and to issue standards in this area in order to make the results more evident.

Soft measurement methods and corresponding instruments will become widespread both in the near and more distant future. To ensure trustworthiness of their results, methods for diminishing the subjectivity role in expert evaluations should be developed and included in international standards.

References

[1]. International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM), 3rd Edition, JCGM 200:2012, *Joint Committee for Guides in Metrology*, 2012.

[2]. MI Recommendation 2021-89, GSI, Metrological Assurance of Flexible Manufacturing Systems, Fundamentals, *Committee on Standardization and Metrology*, 1991 (In Russian).

[3]. K. V. Sapozhnikova, R. Ye. Taimanov, V. V. Kochugurov, Metrological checking as a component of diagnostics of flexible production systems and robotics complexes, in Testing, Checking and Diagnostics of Flexible Production Systems, *Nauka*, Moscow, 1988, pp. 269-273 (in Russian).

[4]. M. P. Henry, D. W. Clarke, The self-validating sensor: Rationale, definitions and examples, *Control Engineering Practice*, Vol. 1, Issue 4, 1993, pp. 585-610.

[5]. R. Taymanov, K. Sapozhnikova, What makes sensor devices and microsystems 'intelligent' or 'smart'?, Chapter 1, in Smart Sensors and MEMS for Industrial Applications (S. Nihtianov, A.L. Estepa, Eds.), 2nd Edition, *Woodhead Publishing, Elsevier Limited*, 2018, pp. 1-22.

[6]. R. Taymanov, A. Pronin, K. Sapozhnikova, Iu. Baksheyeva, I. Danilova. Actual measuring technologies of Industry 4.0 and analysis of their realization experience, *Journal of Physics: Conference Series*, Vol. 1379, 2019, 012049.

[7]. M. P. Henry, Spectral analysis techniques using prism signal processing, *Measurement*, Special Issue, 2021 (in print).

[8]. V. Uspenskiy, K. Vorontsov, V. Tselykh, V. Bunakov, Discrete and fuzzy encoding of the ECG-signal for multidisease diagnostic system, in Advanced Mathematical and Computational Tools in Metrology and Testing X (F. Pavese, W. Bremser, A. Chunovkina, A. B. Forbes, Eds.), *World Scientific Publishing Company*, 2015, pp. 377-384.

[9]. S. Muravyov, S. Tao, M. C. Chan, E. Tarakanov, Consensus rankings in prioritized converge-cast scheme for wireless sensor network, *Ad Hoc Networks*, Vol. 24, Part A, January 2015, pp. 160-171.

[10]. R. Taymanov, K. Sapozhnikova, A. Ionov, Topical metrology problems in the era of cyber-physical systems and Internet of Things, in *Proceedings of the 18th International Congress of Metrology*, 19-21 September 2017, Paris, France, https://cfmetrologie.edpsciences.org/articles/metrology/pdf/2017/01/metrology_metr2017_09006.pdf

[11]. R. Taymanov, K. Sapozhnikova, Trend in efficient development of measuring instruments and systems, in *Proceedings of the XXIX International Scientific Symposium 'Metrology and Metrology Assurance' (MMA'19)*, Sozopol, Bulgaria, 6-10 September 2019, pp. 1-5.

[12]. K. Sapozhnikova, Iu. Baksheeva, R. Taymanov, Features and experience of metrological self-check organisation in multichannel measuring system, in *Proceedings of the XXX International Scientific Symposium 'Metrology and Metrology Assurance' (MMA'20)*, Sozopol, Bulgaria, 7-11 September 2020, pp. 1-5.

[13]. K. E. Harper, C. Ganz, S. Malakuti, Digital twin architecture and standards, *IIC Journal of Innovation*, November 2019, pp. 1-12.

[14]. A. Bécue, E. Maia, L. Feeken, P. Borchers, I. Praça, A new concept of digital twin supporting optimization and resilience of factories of the future, *Applied Sciences*, Vol. 10, Issue 13, 2020, 4482.

[15]. S. V. Prokopchina, Soft measurements: Methodology and application in scientific technical and socio-economic tasks of digital economy, *Soft Measurement and Computing*, Vol. 10, Issue 9, 2018, pp. 3-33 (in Russian).

[16]. L. Mari, P. Carbone, D. Petri, Fundamentals of hard and soft measurement, Chapter 7, in Modern Measurements: Fundamentals and Applications, *Wiley-IEEE Press*, 2015, pp. 203-262.

[17]. GOST R 8.673-2009, State System for Ensuring the Uniformity of Measurements, Intelligent Sensors and Intelligent Measuring Systems, Basic Terms and Definitions, *Standartinform*, 2010 (in Russian).

[18]. GOST R 8.734– 2011, State System for Ensuring the Uniformity of Measurements, Intelligent Sensors and Intelligent Measuring Systems, Methods of Metrological Self-checking, *Standartinform*, 2011 (in Russian).

(038)

Automation of Distributed Computing in a P2P Network

Y. Shichkina [1], M. Kupriyanov [1], K. Krinkin [1] and S. Moldachev [1]
[1] Saint Petersburg Electrotechnical University "LETI", Department of Computer Science and Engineering,
5 prof. Popova street, Saint Petersburg, Russia
E-mail: strange.y@mai.ru

Summary: This article describes new methods for distributing tasks among nodes of a P2P computing network, taking into account the structure of tasks and data. In particular, the article contains cases when the task can be divided into independent parts according to the data and when the subtasks are related. We also present the test results and the minimum requirements required to execute the methods. The methods are based on graph theory and have good performance in finding a resource suitable for solving a problem in a P2P system. The article also describes a method for selecting a target node in a P2P network, which is additional to the basic methods.

Keywords: P2P computing network, Distributed computing, Graph, Set, Optimization.

1. Introduction

At the moment, P2P networks are increasingly used in the modern world to build distributed computing systems (P2P grid). Examples of such systems: BonjourGrid, Condor – Flock P2P, P2PGrid, Alchemi, PastryGrid, Self-Gridron, Harmony.

In P2P computing systems, each user provides computing power to perform tasks that are sent to other network participants. Since the system is highly distributed, it can be difficult to use centralized task distribution methods. Instead, each node of a P2P computing system has its own task distributor that determines whether a task can be completed on the current node or should be sent to another node. At the same time, the task distributor performs dynamic load balancing on the system.

Achieving high performance P2P computing system is a difficult problem due to the heterogeneous nature of the distributed network, where the computing capabilities of each device can be radically different. Various heuristic strategies are known today for solving the problem of heterogeneity.

Classical task distribution methods aim to minimize the execution time of all tasks, but this cannot be applied in a highly distributed environment. This is due to the fact that obtaining all possible parameters of a node can take quite a long time. Taking this problem into account, some researchers have proposed methods for distributing tasks using the fairness index [1-3]. Using this index in load balancing methods, it becomes possible to evenly distribute the load across the entire network.

It is also necessary to mention the Jain fairness index [4], which is widely used to determine the uniform network load. In [5], the author tries to improve system performance by maximizing the fair load index of the entire system.

However, most of the existing solutions that try to eliminate the inequitable distribution of resources are based more on system operating experience and do not always lead to an increase in system performance.

This article proposes our own approach to solving this problem.

2. Formulation of the Problem

Any P2P network can be represented as a directed weighted graph, where the nodes of the network with certain parameters are the vertices of the graph, and the edges between the vertices are the links between the nodes. The edge weight is an indicator of the latency of data transmission between nodes. An example of such a graph is shown in Fig. 1.

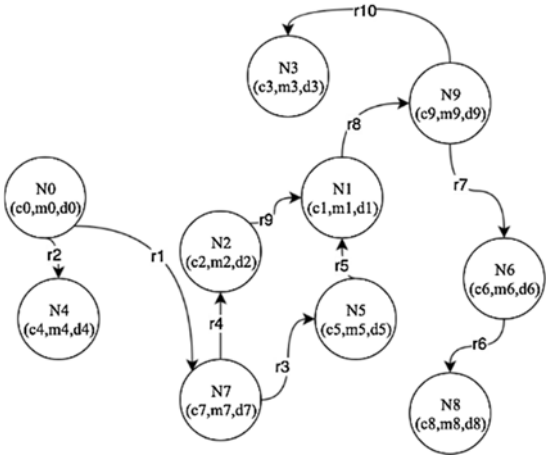


Fig. 1. Example of a weighted graph $G(c_n, m_n, d_n, r_n)$ describing a P2P network, where: c_n – the number of CPU cores, m_n – the amount of free RAM, d_n – the amount of free disk space, a r_n – the delay between nodes.

Suppose there is a weighted directed graph $G(V, R, Pv, Pr, t_0)$, in which V is the set of vertices corresponding to the nodes of the P2P network, R is the set of edges corresponding to connections between nodes, Pv is the set of parameters of nodes, Pr is the set of edge parameters, t_0 is the time instant at which the graph is created.

The task of optimizing the operation of a P2P network is reduced to the task of optimizing routes along a graph in accordance with the given weights and optimization parameters.

3. Methods for Selecting a Target Node and Distribution of Tasks Among the Nodes of a P2P Network

Input data for the method are: graph G, set of weights of edges P_r, set of weights of vertices P_v.

The output data are the numbers of nodes that correspond to a potential executor node and to which there is a minimum path from a given customer node.

The essence of the method for selecting a target node:

1) Find nearby executor nodes: $V' = \left\{v \in G \mid |(v, v_0)| = \min_{v \in G} |(v, v_0)| \,\&\, P_r = \text{true}\right\}$, where $|(v, v_0)| = k$ is the length of the path between the vertices v and v_0;
2) If there are no such nodes, i.e. $V' = \emptyset$, then step 6. Else, if $V \neq \emptyset$, then go to step 3;
3) For each found node $v' \in V'$ it is necessary to check the execution of the set of rules P_v : $V'' = \{v' \in V' \mid P_v = true\}$;
4) If $V'' \neq \emptyset$, then step 6. Otherwise, step 5;
5) Put $v_0 = V'$ and $P_r = P_r - r'$ for all r', corresponding to $v' \in V'$ and step 1;
6) End of method.

The essence of the method for distribution of tasks:

1) Using the method of determining suitable nodes, find nodes from the set $\boldsymbol{V}$ satisfying the conditions $\boldsymbol{P_r}$ and $\boldsymbol{P_v}$. Initialize empty list $\boldsymbol{Rts} = []$;
2) If $\boldsymbol{V'} \neq \emptyset$, then the result of the method is the exception "No resources";
3) For each task $\boldsymbol{ts_i} \in \boldsymbol{TS}$, $\boldsymbol{i} = \boldsymbol{1, 2, \ldots, mt}$ perform the steps:
   a) If the current number of retries is $n < N$, then send the task to the node $v' \in V'$. Otherwise, take the next task and step 3;
   b) Wait for a response from the executor node about the execution of tasks or after the waiting time expires go to step d;
   c) If the response from the executor node has the status "executed with an error" then step d. Otherwise, add the node's response to Rts and go to step 3;
   d) Remove the current vertex v' from V', increase n by 1 and go to step 3.
4) Returns a list of responses $\boldsymbol{Rts}$. The end of the method.

5. Testing the Method of Distribution of Tasks among the Nodes of a P2Pp Network

Testing is performed on a P2P network topology of 3000 nodes. The tasks were launched gradually with a delay of 1-3 ms. The number of tasks is 18000. After every 500th task started, the system calculated the number of rejected tasks and the fairness index for the entire P2P network using the formula [10]:

$$Fairness\ Index = \frac{(\sum x_i)^2}{n \sum {x_i}^2}$$

Below are the results of testing the method of distributing tasks over a P2P network, taking into account various criteria for the required CPU quota (Fig. 2).

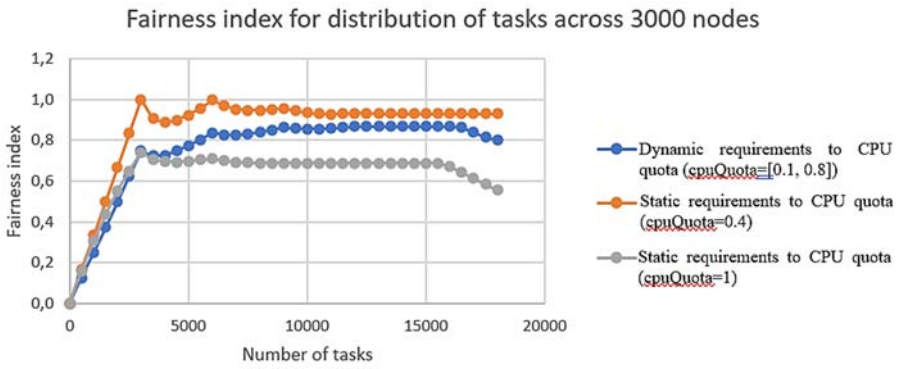


Fig. 2. Fairness index graph depending on the criterion of the required CPU quota.

From Fig. 2 it can be seen that with higher CPU quota requirements, the fairness index is lower, and the number of rejected tasks is higher. This is due to the fact that the method cannot find suitable nodes because there are not enough available resources on the nodes. In this case, the fairness index shows that the P2P network is loaded by 70 %, therefore the remaining 30 % of the resources remain unused.

4. Conclusions

The considered methods can be used in work in various P2P networks, some of them are centralized, others are completely decentralized. Some methods make requirements on the structure of the P2P network, others can work in unstructured highly decentralized networks. The testing shows that the task

distribution method evenly distributes tasks over a P2P network and can be used to solve computational tasks in a P2P network, transfer content and other purposes.

Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement № 075-15-2020-933 dated 13.11.2020 for the implementation of state support for the establishment and development of the world-class scientific center «Pavlov center «Integrative physiology for medicine, high-tech healthcare, and stress-resilience technologies».

References

[1]. T. Repantis, Y. Drougas, V. Kalogeraki, Adaptive component composition and load balancing for distributed stream processing applications, *Peer-to-Peer Networking and Applications*, Vol. 2, Issue 1, 2009, pp. 60-74.

[2]. Y. Drougas, V. Kalogeraki, A fair resource allocation algorithm for peer-to-peer overlays, in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'05)*, Vol. 4, 2005, pp. 2853-2858.

[3]. G. Di Fatta, M. R. Berthold, Decentralized load balancing for highly irregular search problems, *Microprocessor Microsystems*, Vol. 31, 2007, pp. 273-281.

[4]. R. K. Jain, The art of computer systems performance analysis: techniques for experimental design, in Measurement, Simulation and Modelling, *John Wiley & Sons*, 1991.

[5]. T. Repantis, Y. Drougas, V. Kalogeraki, Adaptive resource management in peer-to-peer middleware, in *Proceedings of the Workshop on Parallel and Distributed Real-Time Systems (WPDRTS'05)*, 2005, pp. 132-140.

(039)

Pulse Averaging Primary Converters for Monitoring Systems

O. Bureneva, P. Bondarenko and N. Safyannikov
St. Petersburg State Electrotechnical University 'LETI', Department of Computer Science and Engineering
ul. Professora Popova 5, St. Petersburg, 197376, Russian Federation
Tel.: + 7(812)2342503
E-mail: oibureneva@etu.ru

Summary: During the development of monitoring systems, the tendency of transferring computations to sensors to perform primary transformations at the end points is evident. This allows minimizing the volume of transferred data and simplifying the working of the computational core of the system. At the same time, there is a need to develop hardware converters that can solve calculating issues near the sensor. Obviously, such devices should process signals in the form that is used for representation of the sensor output. In most cases these are analog and pulse signals. The paper proposes a generalized structural diagram of a fault-tolerant averaged converter that simultaneously processes analog and pulse-width modulated signals, that can be used to design converters operating near a sensor, and considers an example of implementation of such a primary converter.

Keywords: Primary converter, Sensor, Monitoring, Pulse streams, PWM signal, Streaming calculations.

1. Introduction

During creation of control and monitoring systems using a large number of sensors an important task is to organize the information transmission and processing. Traditional approaches are based on centralized technology: the information received from sensors is collected and processed in a single computing core.

It is possible to increase the efficiency of processing by decentralizing the computations [1]. Decentralized pulse monitoring system has a tree structure. Sensors are placed at the end points of the monitoring system. They are under the direct influence of the phenomenon, physical object or substance and provide primary measurement information. This structure implies a transfer of calculations near the sensors. This will reduce the amount of data transmitted over communication channels, reduce the load on the central computing core, increase the reliability of the system and minimize energy consumption [2].

The transfer of computational resources to sensitive elements leads to the need to develop special hardware calculators for processing pulsed signals.

2. Pulse Primary Transducers

2.1. Format of Processed Data

The output signal of the sensor is most often in analog or frequency form. Working with pulse signals is due to the fact that the calculators near the sensor must process the signal in the form in which it is formed at the sensor output. If the information on the sensor output has a frequency form (pulse flow or PWM), it is not necessary to change, then the construction of the primary converter does not require changing the form of the signal.

If the sensor output signal is in analog form, it is necessary to convert the result into pulse form. For such conversions, appropriate methods and devices have been developed [3, 4].

Constants and scale factors in such devices can be represented as codes. Processing of pulse and PWM signals (quasi-digital) together with binary codes can be performed using digital element base.

All functional conversions of quasi-digital signals are performed in streaming form using digital elements. The result can be presented both in code and pulse form. The pulse form is preferred, as it provides a more reliable and cost-effective transfer of data to the next level of the monitoring system.

Each subsequent system level integrates data from nodes on the previous level and performs processing also in pulse form. Thus, before the data arrives at the central core, much of the computation will be done.

2.2. Elemental Basis of Transducers

The element base for construction of pulse averaging converters for monitoring systems can be represented by logical primitives, because the simplest element of a pulse stream is a single bit. In addition, the basis should include elements that allow to convert streams of different types into code representations, as well as to perform the simplest arithmetic transformations (e.g., increment/decrement). The following elements can be used as a basis when creating pulse averaging converters: element AND, which implements the function of multiplying the pulse stream by the PWM signal; OR element, which provides summation of pulse streams, provided that the moments of arrival of unit pulses do not coincide; reverse counter, which performs the operation of pulse train subtraction with simultaneous accumulation of difference

in time; binary frequency multiplier, which allows forming pulse streams on the basis of binary code.

Pulse flow conversion units can be constructed using a fault-tolerant tracking structure [5], which provides processing of flow data (streams of single pulses and streams of pulse-width-modulated signals) together with binary codes. Compensation mechanisms in such devices are implemented using negative feedback, which ensures achievement of equilibrium state, which is characterized by dynamic equilibrium of input and compensating streams. In the state of dynamic equilibrium, the parameters of the output stream formed by the device correspond to the result of functional transformations.

3. Averaging Pulse Transducer

Using the proposed approach, a device for determining the average value of some physical quantity shown in Fig. 1 is developed. Different sensors representing the output signal in analog (S_A) and PWM form (S_{PWM}) are used for monitoring.

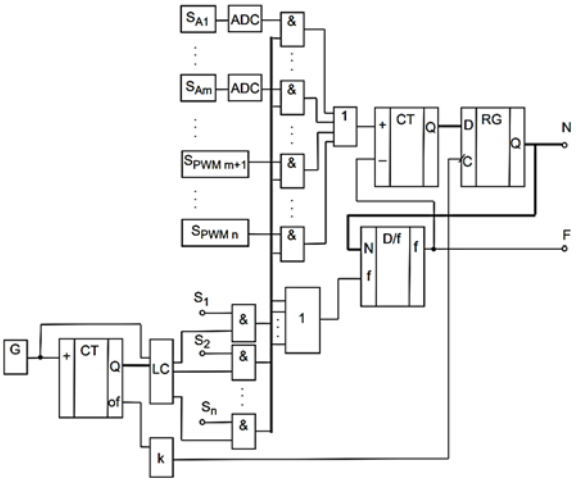


Fig. 1. Averaging pulse transducer.

The device operates in continuous mode and provides measurement of the average value of the measured physical parameter.

The signals from all *n* sensors are read simultaneously. PWM signals from sensors *1* to *m* are used in calculations without conversions. Signals from analog sensors *m+1* to *n* are converted into PWM signals. All PWM signals are summed in the process of their parallel filling with reference frequency from the generator G and averaged. The time averaging is performed by accumulating the measurement results of the previous periods of PWM signals by organizing a tracking pulse system with memory feedback. Averaging over the number of sensors is performed by taking into account only *k* working sensors and excluding the faulty ones by blocking the corresponding channels of forward and feedback signal S_i. In code form the function of the device is as follows

$$N = \frac{1}{k}\sum_{i=1}^{n} S_i N_i,$$

where N_i code equivalent of the sensor output signal. The device output signal in code form is outputted to output N, and in frequency form − to output F.

4. Conclusions

We propose an approach to the organization of pulse averaging converters, providing the processing of signals represented in quasi-digital forms. Transformations are performed in pulse form in the basis of the increment/decrement operation without the use of hardware calculators. The negative feedback provides operation of compensation mechanisms, which provides noise immunity of its operation. The proposed converter is based on simple logic elements, which allows to realize it on FPGA. The universality of the used components allows its production in the form of ASIC.

Acknowledgements

The presented researches have been performed in Saint-Petersburg Electrotechnical University (LETI) during design of the path of primary processing of stream data received from radar. It is supported by the Agreement № 075-11-2019-053 dated 20.11.2019 (Ministry of Science and Higher Education of the Russian Federation, in accordance with the Decree of the Government of the Russian Federation of April 9, 2010 No. 218), project «Creation of a domestic high-tech production of vehicle security systems based on a control mechanism and intelligent sensors, including millimeter radars in the 76-77 GHz range».

References

[1]. C. Bergera, A. Heesa, S. Braunreuthera, G. Reinharta, Characterization of cyber-physical sensor systems, *Procedia CIRP*, Vol. 41, 2016, pp. 638-643.

[2]. M. H. Najafi, S. R. Faraji, K. Bazargan, D. Lilja, Energy-efficient near-sensor convolution using pulsed unary processing, in *Proceedings of IEEE 30th Int. Conference Appl.-Specific Syst. Archit. Processors (ASAP'19)*, New York, USA, 15-17 July 2019, p. 36.

[3]. H. Wang, M. Zhang, Y. Liu, High-resolution digital-to-time converter implemented in an FPGA chip, *Applied Sciences*, Vol. 7, Issue 1, 52.

[4]. N. Samimian, M. Mousazadeh, A. Khoie, A time-based all-digital analog to digital converter for IOT applications, in *Proceedings of 27th Iranian Conference on Electrical Engineering (ICEE'19)*, Yazd, Iran, 30 April – 2 May 2019.

[5]. O. Bureneva, N. Safyannikov, Bit-stream functional converters for decentralized sensor systems, in *Proceedings of 9th Mediterranean Conference on Embedded Computing (MECO'20),* Budva, Montenegro, 8-11 June 2020.

(041)

A New Diagnostic Marker for Endometriosis – Kisspeptin Evaluated in Endometrium with Algorithms of Computer Vision and Machine Learning

A. O. Drobintseva [2], A. S. Krasichkov [1], M. S. Kupriyanov [1] and V. O. Polyakova [1, 2]
[1] Saint Petersburg Electrotechnical University "LETI", 197376, Russia, Saint Petersburg, Professora Popova str., 5,.
[2] Saint Petersburg State Pediatric Medical University, 194100, Russia, Saint Petersburg, Litovskaya str., 2
Tel.: + 79811853607
E-mail: anna.drobintseva@gmail.com

Summary: Current investigation is devoted on processing of medical images obtained with confocal microscope for pathomorphological analysis. The object of the study is endometrium of women suffering from external genital endometriosis, severe persisting desise, lead to infertility. Immuniohistochemical identification of markers such as promising kisspeptin was held in endometrium biopsies from women with different stages of endometriosis. Processing Algorithms of cell nuclei detection and designation of endometrial glands was developed. Also all regions of interest with expression of kisspeptins were marked in confocal images using computer vision at first and mashine learning as more advised method.

Keywords: External genital endometriosis, Optimization of diagnosis, Machine learning, Computer assisted diagnostics, Convolutional Neural Networks.

1. Introduction

The study of the morphology of cells and tissues has always been an important process, both in the field of biology and in the field medicine. Currently, with the existence of many different methods and a wide variety equipment for obtaining experimental morphological data, in the form of digital images, analysis this data is still cumbersome and time consuming. The problem could be solved with using the methods of computer analysis, which are gaining popularity in modern world [1]. Development of molecular markers based on the identification of key nodes of pathogenesis this disease, will create a new diagnostic algorithm, which is necessary to optimize and classification of morph-functional changes, assessing the prognosis of the course of the internal genital endometriosis (IGE) and choosing the optimal targeted therapy [2].

2. Materials and Methods

In immunohistochemical (IHC) study The study included 43 patients with EGE aged from 24 to 36 years old (average age - 33.6 ± 1.3 years). The diagnosis of EGE was established using a laparoscopic scopy. For IHC studies, primary monoclonal antibodies to kisspeptin (1: 140, Abcam, UK). Alexa Fluor 488 (1:1000, Abcam) were taken as secondary antibodies. The nuclei were counterstained with Hoechst33258 dye. Scanning of samples was performed on a microscope FluoView1000 (Olympus). Algorithms of binarization, morphology, segmentation, color correction and contour analysis were used as computer vision algorithms. Convolutional neural network was trained with set of 1000 of confocal images.

3. Results

Studies have shown that the task of detecting markers can be solved with the help of classical computer vision algorithms. However, such algorithms show a detection percentage of markers of the order of 80 %. This is due to the sensitivity of the algorithms to changing the scale and variations in the illumination of the preparation, as well as to the case of superposition of nuclei.

Also we trained neural network to analyze weather it is a nuclei of the cell on image or artefact object. Results are shown in Fig. 1.

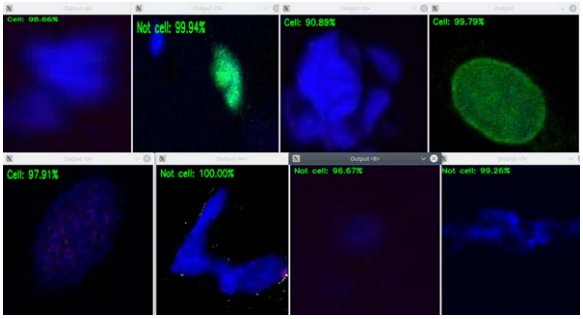


Fig. 1. Sensitivity of neural network in detection of stained with fluorescent dye cell nuclei.

The quality of the algorithm was improved to 90-95 % using machine learning. In contrast to

computer vision algorithms, machine learning algorithms make it possible to identify statistical patterns about the shape of nuclei and the color of markers on a large number of preparations. As a result, the algorithm based on machine learning can be successfully applied not only to images of the drugs that were trained, but also to a whole class of similar drug images.

In terms of correct diagnosis it is important to localize morphological structures in which marker expression is detected. So we developed an algorithm for search of endometrial glands in whole endometrium tissue image. For each image there were internal contours of the glands detected algorithm – detectedContours and internal contours glands marked by the pathologist – labelled. The area of intersection of the marked and of detected contours intersectArea and total the area of the marked outlines – labelledArea.

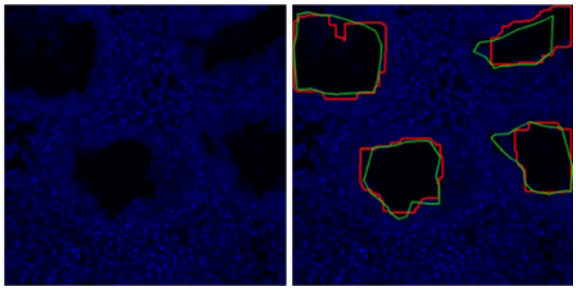


Fig. 2. Detection of glandular structure of endometrium.

The carried out immunohistochemical study revealed that kisspeptin expression in 42 patients women suffering from EGE. One patient did not have identified protein kisspeptin. In our study studied II (n = 19) and III (n = 23) degree IGE, with it was shown that the expression level of KISS1 in patients with a more severe form of the disease was significantly lower than in patients with more mild. Analysis of the obtained data seemed to be more expressive of KISS1 is the cylindrical cells of the endometrial glands, not connective tissue cells stroma. There was a tendency to decrease expression of KISS1 with age.

4. Conclusions

Thus, the creation of an analyzer program using neural networks for the analysis of micrographs and quantitative determination of the expression of diagnostically significant markers will create a new diagnostic algorithm, which is necessary for the optimization and classification of morph-functional changes, assessment of the prognosis of the course of the disease and the choice of optimal targeted therapy.

Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement № 075-15-2020-933 dated 13.11.2020 on the provision of a grant in the form of subsidies from the federal budget for the implementation of state support for the-establishment and development of the world-class scientific center «Pavlov center «Integrative physiology for medicine, high-tech healthcare, and stress-resilience technologies».

References

[1]. M. N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, B. Yener, Histopathological image analysis: A review, *IEEE Reviews in Biomedical Engineering*, Vol. 2, 2009, pp. 147-171.

[2]. A. O. Drobintseva, A. S. Averkieva, M. A. Petrosyan, 3D cultures of endometrial cells: opportunities and prospects, *Tsitologiya*, Vol. 62, Issue 8, 2020, pp. 535-541.

(042)

Age Changes in the Expression Level of Dense Contact Markers in Women after Myomectomy

V. O. Polyakova [1,2], T. S. Kleimenova [2], A. I. Shapovalova [3], D. S. Medvedev [4], A. S. Krasichkov [1] and M. S. Kupriyanov [1]

[1] Saint Petersburg Electrotechnical University "LETI", Prof. Popova, 5. St. Petersburg, 197376, Russia
[2] Saint Petersburg State Pediatric Medical University, Litovskaya str., 2, St. Petersburg, 194100, Russia
[3] Saint-Petersburg Institute of Bioregulation and Gerontology, 3 pr. Dinamo, St. Petersburg 197110, Russia
[4] North-West State medical University named after I.I. Mechanicov, Kirochnaya, 41,
St. Petersburg, 191015, Russia
Tel.: +79062603484
E-mail: kleimenovats@gmail.com

Summary: Uterine fibroids (leiomyomas) are benign tumors that exhibit various forms of smooth muscle differentiation. This is a common pathology that occurs in many women, the number of women with this pathology is growing every year, and so uterine fibroids are a significant socio-economic problem. Claudins (CLDNs) are major tight junction proteins, exhibiting varying tissue expression, with the CLDN expression profile being representative. CLDNs play an important role in neoplastic processes, as they are involved in the formation of a single signaling pathway between the extracellular matrix and the intracellular cytoskeleton.

Keywords: Uterine fibroids, Leiomyomas, Claudins, Uterine scarring, Age-related changes, After myomectomy.

1. Introduction

Uterine fibroids, also called leiomyomas, are the most common benign tumors in women [12, 20]. The impact of fibroids on the economy is very great; about 11 million women in the world suffer from it and their number is steadily growing [14]. Symptoms include menorrhagia, dysmenorrhea, dyspareunia, pelvic pain, and frequent urination [13]. Despite the fact that fibroids are such a common problem, research on etiology is ongoing. Most of the current clinical guidelines indicate that conservative myomectomy has a positive effect on fertility, increases the likelihood of pregnancy and should be performed as a pre-gravid preparation [1]. However, there is also an opposite opinion, its supporters question the advisability of myomectomy in nulliparous women, given the formation of postoperative myometrial scar, which is becoming a more significant complication of pregnancy and childbirth than the fibroid itself [6].

Claudins are small integral membrane proteins that are key tight junctions (PC) molecules that regulate the diffusion of ions and water, and play a role in the formation of cell polarity, adhesion, differentiation, and proliferation [8]. Tight contacts are one of the methods of intercellular adhesion in sheets of epithelial or endothelial cells; they form continuous seals around cells and serve as a physical barrier to prevent the free passage of solutes and water through the intercellular space [23]. PCs exhibit wide variability in density in different organs, ranging from almost complete contraction of the paracellular cleft for solutes to the formation of paracellular pores for certain cations (and anions), for example, in the renal tubules. The permeability of these cell junctions and the function of the epithelial barrier are primarily mediated by claudins, and the isoforms of claudin expressed in the tissue determine the tissue-specific barrier characteristics [5]. Mammalian intracellular claudins contain ~ 7 N-terminal amino acids, ~ 12 loop amino acids, and 25-55 C-terminal amino acids. Claudin density and pore function are highly dynamic. Phosphorylation provides short-term regulation of claudins. CLDN1 is a gene encoding the Claudin 1 protein. Claudin 1 is expressed to a greater extent by the plasma membrane, and the protein can also be found in the cytoplasm, but to a much lesser extent. For a long time, it was considered a putative suppressor of many tumors. Claudin 7 maintains cell polarity, plays an important role in intercellular communication and epithelial cell homeostasis [4]. It is known that Claudin 3, 4, 7 and 8 reduce the permeability of epithelial tissue [7, 18]. Claudin 10 is encoded by the CLDN10 gene and occurs in two isoforms: Claudin 10a and Claudin 10b, which constitute intercellular anion or cation channels, respectively [15]. Claudin 10a and -10b differ significantly not only in their functions, but also in the localization of expression in the body. Expression of Claudin 10a appears to be limited to the kidneys, while Claudin 10b has been found in many tissues: kidney, skin, salivary glands, sweat glands, brain, lungs and pancreas [9, 11, 24].

It is known that with age, the barrier function of tight junctions is impaired. A number of studies have shown that the expression of claudins usually decreases with age in tissues in the liver, lungs [21], kidneys [10, 19] and the pancreas [3], providing a possible mechanism for reducing the barrier function observed in various tissues of the elderly. Parallel studies of human brain micro vessels have revealed an

age-related impairment in the distribution of CLDN5 in postmenopausal women compared with premenopausal women. However, studies of claudins in the tissues of both scars in general and in the scars of intact myometrium have not been described in the literature. From our point of view, the level of expression of claudins will allow us to re-evaluate the viability of the scar and be a prognostic sign for determining the condition of the scar, which is of particular importance for women in older age groups.

2. Aim

The aim of this study was to study the expression level of markers of the functional activity of CLDN1, CLDN7, and CLDN10b tight junctions in biopsy specimens of intact myometrium in women of different age groups.

3. Material and Methods

The examination was carried out in the Department of Operative Gynecology with an operating unit of the FSBSI "The Research Institute of Obstetrics, Gynecology and Reproductology named after D. O. Ott". All patients underwent a set of diagnostic techniques: anamnestic data, clinical and gynecological examination, echography (ultrasound with a transabdominal, transvaginal transducer), endoscopy (hysteroscopy, laparoscopy), histological examination of scrapings and macro preparations removed during operations. Material for immunofluorescence studies was obtained by trephine biopsy of intact myometrium in the area adjacent to the myomatous node during laparoscopic myomectomy in women aged 23 to 47 years. In the control group, myometrium biopsies were taken during diagnostic laparoscopy. The operation was carried out in the 1st phase of the menstrual cycle. All material was divided into 6 groups (Table 1).

Table 1. Study groups.

No.	n	Age	Group
I.	20	23-29	Control
II.	15	30-39	Control
III.	10	40-47	Control
IV.	20	23-29	Patients with uterine fibroids
V.	15	30-39	Patients with uterine fibroids
VI.	10	40-47	Patients with uterine fibroids

Immunofluorescence study was carried out on paraffin sections. Tissue sections with a thickness of 4 μm were placed on glass slides coated with a poly-L-lysine film (Sigma), the preparations were dried for 24 hours in a thermostat at 37 °C. For the immunofluorescence study, a standard two-stage protocol was used with antigen unmasking (high-temperature tissue treatment) in 0.01 M citrate buffer, pH = 6.08-6.10. The following primary monoclonal antibodies were used: Anti-Claudin 1 antibody (rabbit polyclonal, 1: 500, clone: ab15098, abcam), Anti-Claudin 7 / CLDN-7 antibody (rabbit polyclonal, 1: 200, clone: ab27487, abcam), Anti-Claudin 10 antibody (rabbit polyclonal, 1: 300, clone: ab52234, abcam), incubation was carried out in a humid chamber under the conditions specified in the instructions. As secondary antibodies, we used antibodies conjugated with the fluorochrome Alexa Fluor 647 and Alexa Fluor 488 (1: 1000, Abcam); the slides were incubated for 30 min at room temperature in the dark. Cell nuclei were counterstained with Hoechst 33258 (Sigma) for 1 min. The finished preparations were enclosed under cover slips in a Fluorescent Mounting Medium (Dako). As a negative control, a reaction was performed without the use of primary antibodies. The specificity of the antibodies was confirmed in control experiments. The slides were examined using a Zeiss LSM 980 confocal microscope at a magnification of 200×; 5 fields of view were archived from each specimen. c Morphometric analysis was performed using the free software ImageJ, and the relative area of expression, expressed as a percentage, was examined. Statistical processing was carried out in the software Excel 2010. Microsoft Office and in the analytical software Statistica 10.0. Descriptive statistics methods included the assessment of the arithmetic mean (M), the mean error of the mean (m) for features with continuous distribution, as well as the frequency of occurrence of features with discrete values. Using the Mann – Whitney U-test, the samples of indicators of the relative expression area and optical density for all subgroups were compared in pairs. A p-value ≤ 0.05 (5 %) was taken as statistically significant.

4. Results

The study found that the average relative area of the Claudin 1 expression level was statistically significantly lower in all three age groups compared to the control group: 22.5 ± 2.1 % versus 13.7 ± 1.1 %, 18.7 ± 1.7 % - 9.4 ± 0.7 %, 11.2 ± 0.5 % - 3.2 ± 0.2 % ($p \leq 0.05$). Similar data were obtained in the study of Claudin 7: in group I, the mean relative area of expression was 37.3 ± 4.6 %, while in the control group of this group, the increase was 20.5 ± 1.2 %; in group II, this indicator is 26.3 ± 2.9 %, and in group V - 12.4 ± 2.2 %; in group III - 15.2 ± 2.3 %, while in group VI only 6.7 ± 0.6 %, which is also a statistically significant difference ($p \leq 0.05$). The same tendency to decrease was observed in the Claudin 10 marker: in group I - 45.2 ± 1.2 % versus 38.6 ± 2.3 % in the control; 34.6 ± 2.5 % in groups II and 24.2 ± 1.8 % in groups V of women of middle reproductive age; 21.3 ± 1.4 % in III and 15.3 ± 1.3 % in VI studied categories of patients of the older age group ($p \leq 0.05$) (Fig. 1).

All of the above suggests that in patients with uterine fibroids, the level of claudins is statistically significantly lower than in the control group. The results obtained suggest that women from the control group who do not suffer from uterine fibroids can plan pregnancy earlier, in contrast to patients with uterine fibroids, in whom the level of markers of close contacts is much lower. The average age of detecting uterine fibroids is 32 years, so the problem of fibroids and pregnancy is becoming increasingly important due to the rejuvenation of patients on the one hand, and late pregnancy planning on the other [2]. In any case, a pregnant woman after myomectomy should be considered as a patient with an increased risk of complications and a high percentage of the risk of the need for surgical delivery in a planned manner, and, possibly, early delivery. Several studies have shown that during pregnancy there are changes in tight contact proteins such as claudin 1 and 2, as well as changes in the polarity of epithelial cells [17, 25]. Although the function of epithelial differentiation and the function of tight junction proteins in cervical maturation is not fully understood, the changes that are observed in the norm and incorrect expression in preterm labor and post-term pregnancies suggest that these proteins may provide barrier protection during pregnancy [16, 22, 26]. Removal or addition of claudins tends to selectively affect the tight junction barrier function. Certain claudins act as sealants, while others form paracellular ion-selective channels. Thus, claudins can be representative markers for a specific group of patients.

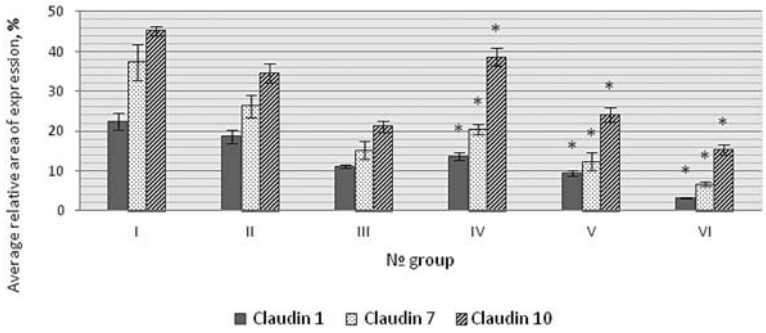


Fig. 1. Diagram of the average relative area of expression of Claudin 1, Claudin 7 and Claudin 10 in the study groups; * $p \leq 0.05$ compared to control.

5. Conclusions

Claudins are a family of proteins that are the most important components of tight junctions, where they establish a paracellular barrier that controls the flow of molecules in the intercellular space between epithelial cells. The level of expression of markers of tight intercellular contacts Claudin 1, Claudin 7 and Claudin 10 is reduced in the epithelium of cell membranes in women with uterine myoma. We assume that a change in the functional activity of tight junctions leads to disruption of the connections between neighboring cells. The expression level of claudins can be used as a marker and target for targeted therapy.

Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement № 075-15-2020-933 dated 13.11.2020 on the provision of a grant in the form of subsidies from the federal budget for the implementation of state support for the-establishment and development of the world-class scientific center «Pavlov center «Integrative physiology for medicine, high-tech healthcare, and stress-resilience technologies».

References

[1]. L. V. Adamyan, E. N. Andreeva, N. V. Artymuk, L. D. Belotserkovtseva, et al., Myoma of the uterus: Diagnosis, treatment and rehabilitation, *Clinical Guidelines for the Management of Patients*, 2015, pp. 164-172.

[2]. E. M. Vikhlyaeva Guidelines for the Diagnosis and Treatment of Uterine Leiomyoma, *MEDpress-inform*, 2004.

[3]. T. D'Souza, C. A. Sherman-Baust, S. Poosala, et al., Age-related changes of claudin expression in mouse liver, kidney, and pancreas, *J. Gerontol. A Biol. Sci. Med. Sci*., Vol. 64, Issue 11, 2009, pp.1146-1153.

[4]. L. Ding, Z. Lu, O. Foreman, et al., Inflammation and disruption of the mucosal architecture in claudin-7-deficient mice, *Gastroenterology*, Vol. 142, Issue 2, 2012, pp. 305-315.

[5]. A. Escudero-Esparza, W. G. Jiang, T. A. Martin, The claudin family and its role in cancer and metastasis, *Front. Biosci*., Vol. 16, 2011, pp.1069-1083.

[6]. S. Fagherazzi, S. Borgato, M. Bertin, A. Vitagliano, L. Tommasi, L. Conte, Pregnancy outcome after laparoscopic myomectomy, *Obstet. Gynecol*., Vol. 41, Issue 4, 2014, pp. 375-379.

[7]. H. Fujita, H. Chiba, H. Yokozaki, N. Sakai, K. Sugimoto, T. Wada, T. Kojima, T. Yamashita, N. Sawada, Differential expression and subcellular localization of claudin-7, -8, -12, -13, and -15 along the mouse intestine, *J. Histochem. Cytochem*., Vol. 54, Issue 8, 2006, pp. 933-944.

[8]. M. Furuse, T. Hirase, M. Itoh, A. Nagafuchi, S. Yonemura, S. Tsukita, S. Tsukita, Occludin: A novel integral membrane protein localizing at tight junctions, *J. Cell Biol*., Vol. 123, 1993, pp. 1777-1788.

[9]. D. Günzel, M. Stuiver, P. J. Kausalya, L. Haisch, S. M. Krug, R. Rosenthal, I. C. Meij, W. Hunziker, M. Fromm, D. Muller, Claudin-10 exists in six alternatively spliced isoforms that exhibit distinct localization and function, *J. Cell. Sci*., Vol. 122, 2009, pp. 1507-1517.

[10]. M. Haddad, F. Lin, V. Dwarakanath, K. Cordes, M. Baum, Developmental changes in proximal tubule tight junction proteins, *Pediatr. Res*., Vol. 57, 2005, pp. 453-457.

[11]. T. Inai, A. Sengoku, X. Guan, E. Hirose, H. Iida, Y. Shibata, Heterogeneity in expression and subcellular localization of tight junction proteins, claudin-10 and -15, examined by RT-PCR and immunofluorescence microscopy, *Arch. Histol. Cytol*., Vol. 68, 2005, pp.349-360.

[12]. B. N. Kashani, G. Centini, S. S. Morelli, G. Weiss, F. Petraglia, Role of medical management for uterine leiomyomas, *Best Pract. Res. Clin. Obstet. Gynaecol*., Vol. 34, 2016, pp. 85-103.

[13]. H. G. Kim, Y. J. Song, Y. J. Na, O. H. Choi, A case of torsion of a subserosal leiomyoma, *Journal of Menopausal Medicine*, Vol. 19, Issue 3, 2013, pp. 147-150.

[14]. E. E. Marsh, A. Al-Hendy, D. Kappus, A. Galitsky, E. A. Stewart, M. Kerolous Burden, Prevalence, and treatment of uterine fibroids: A survey of U. S. women, *J. Womens Health*, 2018, Vol. 27, Issue 11, pp.1359-1367.

[15]. N. Meyers, C. Nelson-Williams, L. Malaga-Dieguez, H. Kaufmann, E. Loring, J. Knight, R. P. Lifton, H. Trachtman, Hypokalemia associated with a claudin 10 mutation: A case report, *Am. J. Kidney Dis*., Vol. 73, 2019, pp. 425-428.

[16]. S. Mittal, Risk of high-grade precancerous lesions and invasive cancers in highrisk HPV-positive women with normal cervix or CIN 1 at baseline-A population-based cohort study, *Int. J. Cancer*, Vol. 140, 2017, Issue 8, pp. 1850-1859.

[17]. L. Peralta, Mechanical assessment of cervical remodelling in pregnancy: Insight from a synthetic model, *J. Biomech*., Vol. 48, 2015, Issue 9, pp. 1557-1565.

[18]. C. Rahner, L. L. Mitic, J. M. Anderson, Heterogeneity in expression and subcellular localization of claudins 2, 3, 4, and 5 in the rat liver, pancreas, and gut, *Gastroenterology*, Vol. 120, 2001, pp. 411-422.

[19]. J. L. Reyes, M. Lamas, D. Martin, M. del Carmen Namorado, S. Islas, J. Luna, M. Tauc, L. Gonzalez-Mariscal, The renal segmental distribution of claudins changes with development, *Kidney Int*., Vol. 62, 2002, pp. 476-487.

[20]. E. A. Stewart, S. K. Laughlin-Tommaso, W. H. Catherino, S. Lalitkumar, D. Gupta, B. Vollenhoven, Uterine fibroids, *Nat. Rev. Dis. Primers*., 2016, Vol. 2, 16043.

[21]. C. G. Tankersley, J. A. Shank, S. E. Flanders, et al., Changes in lung permeability and lung mechanics accompany homeostatic instability in senescent mice, *J. Appl. Physiol*., Vol. 95, 2003, pp. 1681-1687.

[22]. B. C. Timmons, Dynamic changes in the cervical epithelial tight junction complex and differentiation occur during cervical ripening and parturition, *Endocrinology*, Vol. 148, 2007, pp. 1278-1287.

[23]. S. Tsukita, Y. Yamazaki, T. Katsuno, A. Tamura, S. Tsukita, Tight junction-based epithelial microenvironment and cell proliferation, *Oncogene*, Vol. 27, 2008, pp. 6930-6938.

[24]. C. M. van Itallie, S. Rogan, A. Yu, L. S. Vidal, J. Holmes, J. M. Anderson, Two splice variants of claudin-10 in the kidney create paracellular pores with different ion selectivities, *Am. J. Physiol. Ren. Physiol*., 2006, Vol. 291, pp. 1288-1299.

[25]. C. R. Wira, et al., Epithelial cells in the female reproductive tract: a central role as sentinels of immune protection, *Am. J. Reprod. Immunol*., 2005, Vol. 53, pp. 65-76.

[26]. D. Wohlmeister, Association of human papillomavirus and Chlamydia trachomatis with intraepithelial alterations in cervix samples, *Mem. Inst. Oswaldo Cruz*, Vol. 111, 2016, Issue 2, pp. 106-107.

(043)

Modern Methods for Determining Emotional Stress Based on Physiological Signals and Machine Learning

E. Pustozerov [1, 2] and R. Uvarov [1]
[1] Saint Petersburg Electrotechnical University, Professora Popova, 5, 197376 Saint Petersburg, Russia
[2] Almazov National Medical Research Centre, Akkuratova 2, 191014 Saint Petersburg, Russia
Tel.: + 79213684589
E-mail: eapustozerov@etu.ru

Summary: The purpose of the study is to review emotional stress and psychological distress detection techniques based on automatic analysis of physiological signals and machine learning methods, which can be utilized in remote diagnostics. Various methods for assessing the psycho-emotional state of a person are considered, including questionnaires and various types of physiological signals. The most informative signals and signal features for determining stress are highlighted. A comparison of various studies using machine learning algorithms to determine stress is made and appropriate advantages and disadvantages of these methods are highlighted.

Keywords: Emotional stress, Distress, Physiological signals, Deep learning, Monitoring.

1. Introduction

Stress is one of the most important factors negatively affecting human health [1]. A timely detection of emotional and mental stress in convenient ways that do not require complex devices that interfere with human work is necessary. In the review we consider studies devoted to the identification of signals reflecting the level of stress and the most compact systems to assess its level in a quantitative manner.

2. Methods for Assessing Emotional and Mental Stress Based on Questionnaires

There are a number of methods for determining the degree of emotional and mental stress that can be applied during remote monitoring. The most common method of assessing the condition is questionnaires. The Profile of Mood State (POMS) [2] is a well-established measure of psychological stress derived from factor analysis, and its high level of reliability and validity has been documented.

Another way to assess the impact of physiological stress on the body is to analyze the quality of sleep. The Pittsburgh Sleep Quality Index (PSQI) [3] is a widely used method for evaluating changes in subjective sleep quality in both clinical practice and research.

The considered methods for assessing emotional and psychological stress are primarily based on a subjective assessment of the subject's condition.

3. Methods Based on Physiological Signals Analysis

The autonomic nervous system increases heart rate and blood pressure [4] as a response to stress. It was shown [5] that the intensity of the α-rhythm decreases on EEG, while the intensity of the β-rhythm [6] and the level of cortisol increase. The galvanic skin reaction is also a well-known indicator of stress levels [7]. Altogether, the signals of ECG, EEG, FPG, EMG and skin-galvanic reaction can be used for instrumental assessment of the stress level.

The study [8] utilized the Emotiv Epoc+ device, which includes a 14-channel EEG recorder with an accelerometer and an ECG recording module for recording heart rate. Statistical analysis of the converted data showed statistically significant differences between the periods of activity and rest of the subject, which indicates the applicability of these signals for instrumental detection of stress levels.

The study [9] analyzed the stress in response to a pain that occurs at a random moment in time. An EMG signal measured by the trapezius muscle was utilized. The study showed that the stressor can be both cognitive and physical, with the average EMG being highest in the first phase.

4. Methods Based on Machine Learning Algorithms

The study [10] used the following set of signals: ECG with two electrodes; skin conductivity measured with two electrodes; thorax and ambient temperatures; spirogram; acceleration of the movement of the chest, obtained with an accelerometer. J48 decision tree; J48 decision tree using AdaBoost; support vector machines were used to a set of extracted features. An accuracy of 86 % was achieved for ECG and spirogram, and of 92 % for normalized ECG and spirogram. Data from surveys of subjects were taken as ground truth. The resulted classifier that can work without preliminary calibration.

A driver's glove with built-in gyroscope, accelerometer and magnetometer that detects the level

of stress is described in the article [11]. With 22 selected features it was possible to achieve a model accuracy of 94 %. From this study, it can be concluded that the data obtained from the driver's hand acceleration, voice tone and facial image can be used to determine the level of stress.

The aim of the study [12] was to develop an easy-to-use real-time method for detecting stress levels. An infrared camera was used to capture the face, as it gives the clearest image of a person's face in the background light. The system consisted of a face detection module and a facial emotion detection module that included a pre-trained classifier. The model showed an 85 % accuracy when tested in a car.

Algorithm for measuring stress levels using ECG, EMG from the trapezius muscle and HRV was presented in [13]. To obtain features from the signals, the components of the wavelet transform were extracted. The k-nearest neighbors algorithm was used as a classifier. An accuracy of 92 % for determining the level of stress on a five-level scale was achieved.

5. Conclusions

From the articles described above, it follows that the signals of EMG, ECG, EEG, spirogram, acceleration of limb movement, galvanic skin response and facial expressions can be used as non-invasively captured signals for indirectly determining the level of stress. The most accurate methods were based on EMG of the trapezius muscle and EEG, but these measurements require electrodes, which are not suitable in intensive work conditions. Methods based on measuring limb accelerations and analyzing face images provided a high accuracy, but the developed models only performed a binary classification – the absence or presence of stress. Since the magnitude of stress in most experiments was determined through the perceived readings of subjects, the determination of the model's accuracy was based on them.

Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement № 075-15-2020-933 dated 13.11.2020 on the provision of a grant in the form of subsidies from the federal budget for the implementation of state support for the establishment and development of the world-class scientific center «Pavlov center «Integrative physiology for medicine, high-tech healthcare, and stress-resilience technologies».

References

[1]. R. W. Picard, Automating the recognition of stress and emotion: From lab to real-world impact, *IEEE MultiMedia*, Vol. 23, Issue 3, July 2016, pp. 3-7.

[2]. S. Bent, A. Padula, D. Moore, M. Patterson, W. Mehling, Valerian for sleep: A systematic review and meta-analysis, *American Journal of Medicine*, Vol. 119, Issue 12, December 2006, pp. 1005-1012.

[3]. D. J. Buysse, S. Ancoli-Israel, J. D. Edinger, K. L. Lichstein, C. M. Morin, Recommendations for a standard research assessment of insomnia, *Sleep*, Vol. 29, Issue 9, September 2006, pp. 1155-1173.

[4]. Y. M. Ulrich-Lai, J. P. Herman, Neural regulation of endocrine and autonomic stress responses, *Nature Reviews Neuroscience*, Vol. 10, Issue 6, June 2009, pp. 397-409.

[5]. A. A. Nicholson, *et al.*, A randomized, controlled trial of alpha-rhythm EEG neurofeedback in posttraumatic stress disorder: A preliminary investigation showing evidence of decreased PTSD symptoms and restored default mode and salience network connectivity using fMRI, *NeuroImage: Clinical*, Vol. 28, January 2020, 102490.

[6]. J. Bakker, M. Pechenizkiy, N. Sidorova, What's your current stress level? Detection of stress patterns from GSR sensor data, in *Proceedings of the IEEE 11th International Conference on Data Mining Workshops (ICDMW'11)*, December 2011, pp. 573-580.

[7]. A. Liapis, C. Katsanos, D. Sotiropoulos, M. Xenos, N. Karousos, Stress recognition in human-computer interaction using physiological and self-reported data: a study of gender differences, in *Proceedings of the 19th Panhellenic Conference on Informatics*, New York, NY, USA, October 2015, pp. 323-328.

[8]. V. Borisov, A. Syskov, V. Tetervak, V. Kublanov, Mobile brain-computer interface application for mental status evaluation, in *Proceedings of the International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON'17)*, September 2017, pp. 550-555.

[9]. R. Luijcks, H. J. Hermens, L. Bodar, C. J. Vossen, J. van Os, R. Lousberg, Experimentally induced stress validated by EMG activity, *PLoS ONE*, Vol. 9, Issue 4, April 2014, e95215.

[10]. K. Plarre, *et al.*, Continuous inference of psychological stress from sensory measurements collected in the natural environment, in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN'11)*, April 2011, pp. 97-108.

[11]. B. Lee, W. Chung, Wearable glove-type driver stress detection using a motion sensor, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, Issue 7, July 2017, pp. 1835-1844.

[12]. H. Gao, A. Yüce, J. Thiran, Detecting emotional stress from facial expressions for driving safety, in *Proceedings of the IEEE International Conference on Image Processing (ICIP'14)*, October 2014, pp. 5961-5965.

[13]. D. E. Ugarte, D. Linares, G. Kemper, C. A. Almenara, An algorithm to measure the stress level from EEG, EMG and HRV signals, in *Proceedings of the International Conference on Information Systems and Computer Science (INCISCOS'19)*, November 2019, pp. 346-353.

(044)

Intelligent Measurement Technologies for Water Supplying Systems Management

S. Prokopchina
Financial University under the Government of the Russian Federation, Leningradsky pr., 49, Moscow, Russia
Tel.: + 79150070489
E-mail: svprokopchinaMe@mail.ru

Summary: The article proposes the methodology and architecture of the platform for the rapid development of applied intelligent control systems for industrial complexes in conditions of uncertainty. The methodology is based on the regularizing Bayesian approach and technologies based on it. The advantages of this development, consisting in the possibility of integrating artificial intelligence technologies and measurement systems, are considered. An example of a water supply management system is given.

Keywords: Bayesian intelligent technologies, Water supply systems, Measurement.

1. Introduction

The effectiveness of the implementation of the principles of Industry 4.0 largely depends on the intellectualization of methods and tools for solving applied problems, which makes it possible to repeatedly increase the practical usefulness, cognitiveness and speed of problem solving.

The methodology and practice of solving applied problems of Industry 4.0 are related to the implementation of the stages of measurement, transmission and analytical processing of measurement information.

Thus, we can distinguish 3 main blocks of a single information process that provides solutions. All three blocks, as part of a single whole process of solving the problem, should be integrated into the general information space of the solution formation. This means that technologies for implementing the processes of measuring, collecting, transmitting, storing and transforming information based on artificial intelligence methods should be integrated.

It is very important to note that the specifics of the task Industry 4.0 is a significant informational and situational uncertainty expressed in the absence of fully adequate to real objects, models, completeness and accuracy of information, changes in the conditions of realization of information processes and dynamics influencing environmental factors.

Currently, intelligent methods and tools are not integrated with measurement methodologies and systems. This fact causes a number of difficulties in their practical application. On the one hand, artificial intelligence methods and technologies ((neural networks, fuzzy systems, intelligent data processing systems) do not have systems for metrological justification of the information flows of data that they process. And the uncertainty (incompleteness, inaccuracy, vagueness, unreliability, inconsistency) of information determines the instability, incorrectness and unreliability of the obtained solutions. Such solutions cannot be used in practice.

On the other hand, measurement systems that do not have the intellectual capacity do not provide interpretation of the resulting solutions, mainly producing only some pre-processing of measurement data. In addition, with classical measurement schemes, it is not possible to use information in the form of knowledge in the measurement process.

As an example of integration of measurement methods and artificial intelligence, the article considers a hybrid water supply management system based on Bayesian intelligent measurements and technologies [1-3].

2. Methodological Aspects of Creating Intelligent Control Systems for Industrial Complexes Based on Bayesian Intelligent Technologies

Based on the methodology of the regularizing Bayesian approach [1-3] and the technologies of Bayesian intelligent measurements (BIM) and Bayesian integrating technologies (BIT) created on its basis, a platform has been developed for the rapid construction of intelligent control systems for industrial complexes in the conditions of information and situational uncertainty. Currently, intelligent distributed systems for various applications are developed and used in practice. Such systems, inheriting the basic principles of BII and BIT (namely, the integration of different types of data and knowledge flows, metrological justification and the ability to manage the quality of solutions, flexibility and self-developability in the process of functioning), together with promising information technologies for network transmission, collection and intelligent processing of information, represent a new type of systems called Bayesian intelligent networks (BIN).

The need to develop BIN was due on the one hand the requirements of practical problems (usually solved in the conditions of considerable a priori uncertainty) to summarize and use all the available amount of information, on the other hand, features which received the active development of modern network technologies to collect and process servers large volumes of heterogeneous distributed information resources remotely. Unification of information processing in the network is achieved by using accepted standards and languages (e.g. XML), and also technologies of distributed processing (for example, technologies of "cloud" computing, etc.). In these systems, the possibility of convolution of different types of information (data and knowledge) is implemented, its integration into a single information flow that carries the most reliable and complete knowledge about the object or its properties in specific conditions. Such properties of the systems allow us to use all the available information in obtaining solutions, as well as to use processing methods that require sufficient samples in length. All stages of receiving and processing information have metrological support. All solutions are accompanied by a set of metrological characteristics, including indicators of accuracy, reliability and reliability.

The metrology of the Bayesian integrating network makes it possible, in the course of the network operation, not only to identify "bottlenecks" where there is an influx of information with significant distortions and noise, but also to adjust the technology and the list of sources of information resources, and plan an information experiment.

Intelligent GIS networks and systems with analytics are built on the information basis of BIT: Intelligent GIS, which, in addition to the above properties, also have the ability to spatial orientation of decisions, conclusions and recommendations as attributes.

3. The Architecture of the Bayesian Intelligent Platform for Creating Control Systems for Industrial Complexes in Conditions of Uncertainty

Bayesian intelligent platform (BIP) for creating distributed control systems for industrial complexes.- this is a software and hardware complex implemented as an environment for creating user applications based on Bayesian intelligent technologies. The architecture of the BIP is illustrated in Fig. 1.

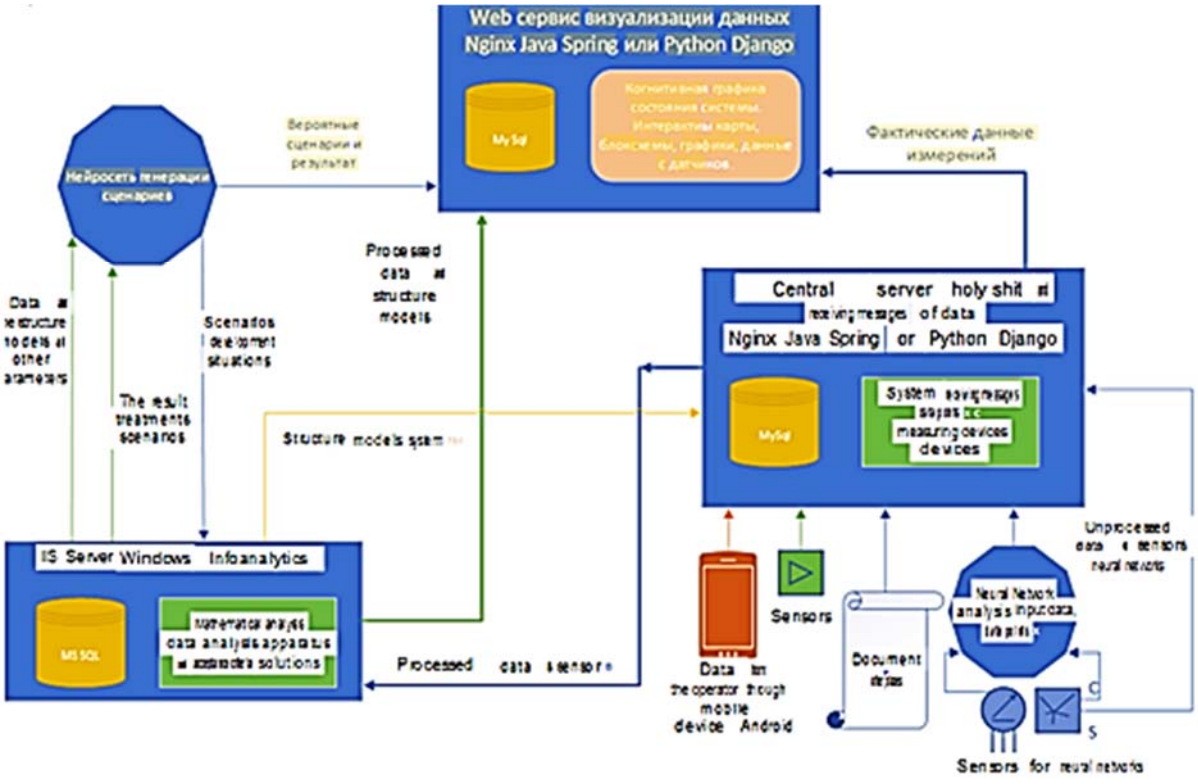


Fig. 1. Architecture of the Bayesian intelligent platform.

BIP includes a system for creating models of complex objects, subsystems, instrumentation and the collection, storage and transfer of data and knowledge (ETL), a system of intelligent data processing on the basis of Bayesian intelligent technologies and neural networks (neural networks and analytical processing of images and texts), the library of Bayesian statistics and applied mathematical methods, models, visual display

of entered and processed information, the system of discharge data and analytical information subsystem of the GIS work, a subsystem of the WEB services.

The main analytical block is implemented on the basis of Bayesian intelligent technologies that allow for the convolution of different types of information in the form of measurement data, specialist knowledge, documentary and video information, and analytical solutions.

The integration function is carried out on the basis of the use of special scales, called conjugate scales with dynamic constraints [2].

The platform allows you to integrate various technologies and implement their methodological principles, in particular, IoT, DATA MINING, BIG DATA, BI, DATA SCIENCE technologies, create monitoring systems, audit, and management decision support.

4. Intelligent Water Supply Management System Based on BIP

The architecture of such an application system corresponds to the architecture of the BIP, shown in Fig. 1, in which all the blocks are focused on the specifics of the application problem. The set of measuring devices is selected in accordance with the measured parameters of hot and cold water supply systems. Neural networks of image processing work with thermal imaging information.

The neural network texts are processing the documentary information. The main intelligent processing of information from measuring instruments, databases of water supply system status logs, neural networks, WEB sources and its integration is carried out on the basis of Bayesian intelligent technologies.

On the basis of this system, it is possible to create production-purpose complexes that allow solving the tasks of internal energy audit and water management audit, accounting for energy consumption and hot and cold water, ensuring energy security of enterprises and territories, training highly qualified specialists in energy management and water supply complexes [4, 5].

5. Conclusion

In conclusion, it should be noted that the proposed methodology and platform for developing application systems enable the integration of artificial intelligence technologies (neural network technology) with the technology of metering systems that allows you to combine the power of systems of receiving and processing a variety of information, including archival information and expertise. The development of various systems of production and socio-economic types does not require the involvement of programmers, it can be quite simply implemented by specialists in the applied field.

Reference

[1]. S. V. Prokopchina, Bayesian intelligent technologies as a methodological basis for big data processing in the conditions of approach uncertainty, in Economics and Management: Problems, Solutions, *Publishing House "Scientific Library"*, Vol. 11, 2019, pp. 105-109.

[2]. S. V. Prokopchina, Principles and methodological aspects of constructing a scale with dynamic constraints for measurements under uncertainty, in Soft Measurements and Calculations, *Publishing House "Scientific Library"*, Vol. 3, 2018, p. 4-15.

[3]. S. V. Prokopchina, Soft measurements and management of complex systems based on the regularizing Bayesian approach, in Economics and Management: Problems, Solutions, *Publishing House "Scientific Library"*, Vol. 5, 2015, pp. 16-25.

[4]. S. V. Prokopchina, Creation of intelligent networks of power engineering and housing and communal services in *Proceedings of the International Conference on Soft Computing and Measurement (SCM'13)*, Vol. 1, St. Petersburg, May 23-25, 2013, pp. 23-27.

[5]. S. V. Prokopchina, A. N. Vetrov, S. Sheremetiev, Automation of water supply enterprise management based on the process approach, *Software Products and Systems*, Issue 1, 2014, pp. 201-205.


9 788409 275380