APPLIED COMPUTER SCIENCE

The Journal is a peer-reviewed, international, multidisciplinary journal covering a broad spectrum of topics of computer application in production engineering, technology, management and economy.

The main purpose of Applied Computer Science is to publish the results of cutting-edge research advancing the concepts, theories and implementation of novel solutions in computer technology. Papers presenting original research results related to applications of computer technology in production engineering, management, economy and technology are welcomed.

We welcome original papers written in English. The Journal also publishes technical briefs, discussions of previously published papers, book reviews, and editorials. Especially we welcome papers which deals with the problem of computer applications in such areas as:

- manufacturing,
- engineering,
- technology,
- designing,
- organization,
- management,
- economics,
- innovations,
- competitiveness,
- quality and costs.

The Journal is published quarterly and is indexed in: BazTech, Cabell's Directory, CNKI Scholar (China National Knowledge Infrastucture), ERIH PLUS, Index Copernicus, J-Gate, Google Scholar, TEMA Technik und Management.

Letters to the Editor-in-Chief or Editorial Secretary are highly encouraged.

CONTENTS

Applied Computer Science, vol. 14, no. 4, pp. 5–11
doi:10.23743/acs-2018-25

Submitted: 2018-06-03
Revised: 2018-09-20
Accepted: 2018-10-15

BOVW, classification, codebook

Baldemar ZURITA [0000-0002-1443-8260]*,
Luís LUNA *, *José HERNÁNDEZ* *, *Federico RAMÍREZ* **

BOVW FOR CLASSIFICATION IN GEOMETRICS SHAPES

Abstract

The classification of forms is a process used in various areas, to perform a classification based on the manipulation of shape contours it is necessary to extract certain common characteristics, it is proposed to use the bag of visual words model, this method consists of three phases: detection and extraction of characteristics, representation of the image and finally the classification. In the first phase of detection and extraction the SIFT and SURF methods will be used, later in the second phase a dictionary of words will be created through a process of clustering using K-means, EM, K-means in combination with EM, finally in the Classification will be compared algorithms of SVM, Bayes, KNN, RF, DT, AdaBoost, NN, to determine the performance and accuracy of the proposed method.

1. INTRODUCTION

The classification of forms is an intriguing and challenging problem found in the intersection of computer vision, geometry processing and machine learning (Li, & Ben Hamza, 2014; Ben Hamza, 2016; Ye & Yu, 2016). The form is an intrinsic characteristic for the understanding of the image, which is stable to illumination and variations in the color and texture of the object. Due to these advantages, the form is widely considered for object recognition (Shaban, Rabiee, Farajtabar & Ghazvininejad, 2013; Wang, 2014; Jia, Fan, Liu, Li, Luo & Guo, 2016). The contours of a form are main characteristics and of great importance for their classification, from these characteristics we can describe the form.

* Apizaco Technological Institute, Computer Systems Department, Apizaco, Tlaxcala, Mexico, e-mail: baldemar.zurita@gmail.com
** Apizaco Technological Institute, Departament of Computer and Systems, Apizaco, Tlaxcala, Mexico

The Bag of Visual Words (BOVW) model proposed by Szelinski (Szelinski, 2011), is based initially on the work of Sivic and Zisserman (2003) as Bag of Words for natural language processing. This method is widely used in classification of objects with excellent results, the main idea of this method is to obtain the description of the images of a training set to generate a codebook or book of visual words, later a classification algorithm can decide the class to which it belongs.

2. CLASSIFICATION BY BOVW METHOD

A classification system that uses the BOVW method has the following phases:

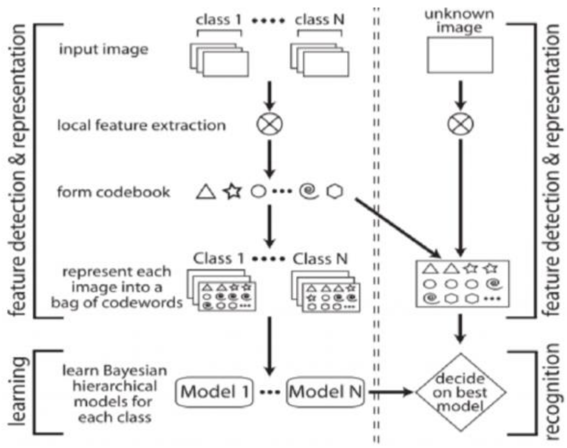


Fig. 1. BOVW model

A classification system that uses the BOVW method has the following phases:
1 – Collect a data set of examples.
2 – Partition the data set into a training set, and a cross validation set (80%–20%).
3 – Find key points in each image, using SIFT.
4 – Take a patch around each key point, and calculate it's Histogram of Oriented Gradients (HoG). Gather all these features.
5 – Build a visual vocabulary by finding representatives of the gathered features (quantization). This done by k-means clustering.
6 – Find the distribution of the vocabulary in each image in the training set. This is done by a histogram with a bin for each vocabulary word.
7 – Train an SVM on the resulting histograms (each histogram is a feature vector, with a label).
8 – Test the classifier on the cross validation set.

If results are not satisfactory, repeat 5 for a different vocabulary size and a different SVM parameters or different classifier.

3. METHOD

We will work with one sets of training, it was obtained from 50 individuals tracing different geometric shapes, 210 images, using 90% for training and 10% for testing, is divided into three classes: triangle, circle and square.

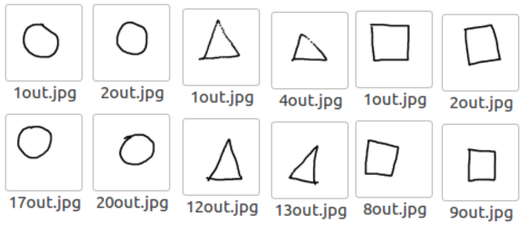


Fig. 2. Dataset of geometric figures

For the modeling of BOVW, a detector and descriptor SIFT and SURF with grouping by K-means, EM and its combination of both K-means + EM classifiers will be used, for the classification step seven algorithms were chosen: vector support machines (SVM) - Supor Vector Machine), random forests (RF – random forest), decision trees (DT – desicion tree), adaptive increase (AdaBoost – adaptive boosting), Bayesian (Gaussian NB), neural network (Neural Network) and nearest neighbors (KNN – k-nearest neighbors) In order to evaluate the performance of each model, each method will be executed 6 times with different values of k (50, 100, 200, 300, 400, 500) in order to obtain the average precision of classification followed by the generation of performance graphs (precision and k-value), cross-validation and confusion matrix analysis will be used to determine the best geometric shape classification model.

All experiments will be executed in a linux operating system, CPU 2.5 Quadcore 64bits, 8 GB of RAM.

4. RESULTS

In the following graphs, the performance of each model with detector and descriptor SIFT and SURF with Clustering Kmeans, EM and Kmeans + EM is shown. Each model was executed with different values of k (dictionary size).

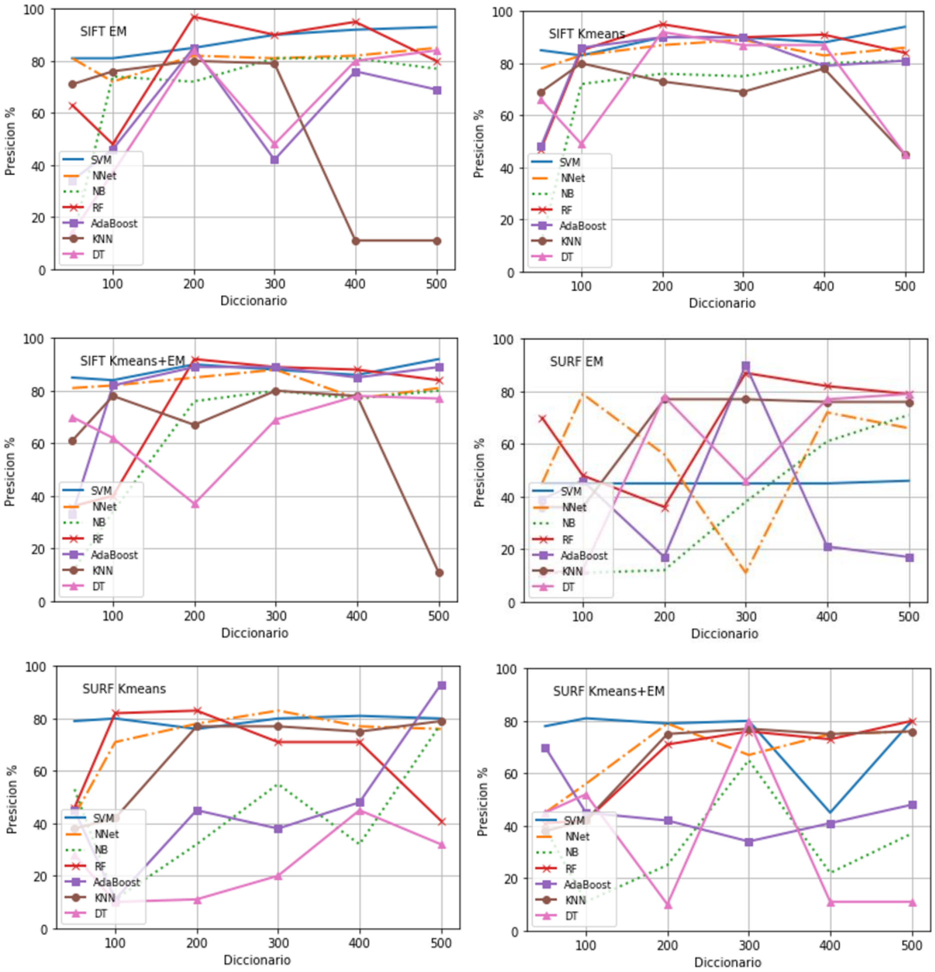


Fig. 3. Comparative chart of the performance of each model

It can be seen that the SIFT descriptor and detector is more stable with the clustering algorithms EM, Kmeans and Kmeans + GM, since a better precision is obtained at the time of the classification for each value of k, in appearance, being gradually more accurate to as the dictionary size increases. Obtaining the values of the cross validation (accuracy, recall, f1 score) can have a better decision on which is the best combination of descriptor, grouping algorithm and classifier, the following comparative table was generated choosing the models with greater accuracy than 90%.

Tab. 1. Results

Detector Cluster	Classifier	K	Precision	Recall	F1-score	Matrix
SIFT EM	RF	200	97	97	97	[20 0 0] [1 19 0] [0 1 19]
SIFT Kmeans	RF	200	95	95	95	[19 1 0] [0 19 1] [0 1 19]
SIFT Kmeans	SVM	500	94	93	93	[18 0 2] [0 18 2] [0 0 20]
SURF Kmeans	AdaBoost	500	93	92	92	[19 1 0] [0 20 0] [0 4 16]
SIFT Kmeans+EM	RF	200	92	90	90	[15 5 0] [0 20 0] [0 1 19]
SIFT Kmeans+EM	SVM	500	92	92	92	[18 1 1] [0 18 2] [0 1 19]
SIFT Kmeans	DT	200	92	92	92	[20 0 0] [0 17 3] [0 2 18]
SIFT EM	SVM	400	92	92	92	[18 0 2] [0 18 2] [0 1 19]
SIFT Kmeans	RF	400	91	90	90	[20 0 0] [3 17 0] [0 3 17]
SIFT Kmeans+EM	SVM	200	90	88	88	[15 3 2] [0 18 2] [0 0 20]
SIFT Kmeans	SVM	200	90	88	88	[14 4 2] [0 19 1] [0 0 20]
SIFT EM	SVM	300	90	88	88	[15 2 3] [0 18 2] [0 0 20]

In this table we can see the best classification results by means of the confusion matrix, which is a tool that allows to visualize the level of precision of a classifier, the rows of the matrix represent the images of the evaluation set (ground-truth) , each row is a different class: circle, square and triangle in a descending order, in the columns we have the same order of the classes from left to right.

Accuracy is measured by calculating the sum of the diagonal of the matrix, which represents the correctly classified images, among the total number of images in the matrix. This table shows the average of the accuracy of the three classes that represents the quality of the classifier response. Precision = TP / TP + FP.

The sensitivity (recall) measures the efficiency in the classification of all elements of the class by means of the calculation of the real positives between the sum of the real positives and the false positives. The average of the sensitivity of the three classes is shown in the table. Recall = TP / TP + FN.

The F1 score can be interpreted as a weighted average of the precision and sensitivity, where a score F1 reaches its best value at 1 and the worst score at 0. F1 Score = 2 * (Recall * Precision) / (Recall + Precision) .

The matrix shows us in its diagonal, the number of correctly classified images. In this case, the model that proved to have the best accuracy is the SIFT detector combination, EM grouping algorithm, k = 200 (200 word dictionary) and the Random Forest classifier. For this model, an additional validation method was added: cross validation.

SIFT EM k = 200 Random Forest
Classification Report

	precision	recall	f1-score	support	Confusion Matrix
Circle	0.95	1.00	0.98	20	[20 0 0]
Square	0.95	0.95	0.95	20	[1 19 0]
Triangle	1.00	0.95	0.97	20	[0 1 19]
avg / total	0.97	0.97	0.97	60	Accuracy Matrix: 96.6 %
					Cross Validation RF: 95.0 %

Fig. 3. Results of the model with better precision with cross validation

Cross-validation is a technique used to evaluate the results of a statistical analysis and ensure that they are independent of the partition between training and test data.

5. CONCLUSION

The classification of images with the BOVW model is fascinating, since it shows an acceptable precision in its labels, in this case, only a dictionary size of 200 words (k = 200) was necessary to be able to classify with accuracy greater than 95% the dataset of geometric shapes exposed in this project, which leads to a lower computational effort, shorter processing time, smaller size in Kb of the dictionary and a faster response by the classifier.

Because the dataset does not contain much information, only a shape descriptor was used, there are other descriptors such as color, which in combination make a more robust system, even classifying more complex images, such as training the algorithm that can discriminate different types of roses for their color, since the shape would be the same, but its variation in color makes it different from the others and therefore, it is more appropriate to use the description of the color. It can be inferred that the success of this model is based on the objects to be classified, since each phase of the method has to be adjusted, changing descriptors, algorithms, classifiers, until obtaining the best result, by means of evaluation methods.

As future work, you could use other grouping algorithms, such as BIRCH or Fuzzy Kmeans, with the aim of raising the accuracy rate to 99%, as well as increasing the size of the dataset and testing this model with other datasets used in computer vision like Caltech101.

REFERENCES

Ben Hamza, A. (2016). A graph-theoretic approach to 3D shape classification. *Neurocomputing, 211*, 11–21.

Jia, Q., Fan, X., Liu, Y., Li, H., Luo, Z., & Guo, H. (2016). Hierarchical projective invariant contexts for shape recognition. *Pattern Recognition*, *52*, 358–374. doi:10.1016/J.PATCOG. 2015.11.003

Li, C., & Ben Hamza, A. (2014). Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: a comparative survey. *Multimedia Systems*, *20*(3), 253–281. doi:10.1007/s00530-013-0318-0

Shaban, A., Rabiee, H., Farajtabar, M., & Ghazvininejad M. (2013). From local similarity to global coding; an application to image classification. In: *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2794–2801). Portland, USA: IEEE.

Sivic, J., & Zisserman, A. (2003). Video Google: a text retrieval approach to object matching in videos. In: *Proceedings of the Ninth IEEE International Conference on Computer Vision – Volume 2* (pp. 1–9). USA: IEEE Computer Society Washington.

Szelinski, R. (2011). *Computer Vision: Algorithms and Applications* (pp. 658–729). Springer Verlag.

Wang, X., Feng, B., Bai, X., Liu, W., & Latecki, L. J. (2014). Bag of contour fragments for robust shape classification. *Pattern Recognition*, *47*(6), 2116–2125.

Ye, J., & Yu, Y. (2016). A fast modal space transform for robust nonrigid shape retrieval. *The Visual Computer*, *32*(5), 553–568. doi:10.1007/s00371-015-1071-5

Applied Computer Science, vol. 14, no. 4, pp. 12–24
doi:10.23743/acs-2018-26

Submitted: 2018-11-08
Revised: 2018-12-12
Accepted: 2018-12-18

economy and management, economy in transport, technological systems

Łukasz WOJCIECHOWSKI [0000-0002-2773-6064]*,
Tadeusz CISOWSKI [0000-0002-9999-4863]**

MODEL OF A COMPUTER SYSTEM FOR SELECTION OF OPERATING PARAMETERS FOR TRANSPORT VEHICLES IN THE ASPECT OF THEIR DURABILITY

Abstract

The article presents a new method of selecting operating parameters of vehicles which allows to prolong their useful life. The theoretical and practical aspects of the new approach were discussed. The key methods of optimizing the use of vehicles and their transport routes were analysed. The proposed method of selecting operating parameters of transport vehicles was based on elements of the dynamic programming model using heuristic algorithms. A model and then an algorithm and a computer program for selecting operating parameters for vehicles were developed.

1. INTRODUCTION

Enormous progress and changes in the operation of enterprises occur in two basic ways: one way is the application of modern technology and technique, the other one is the development of the methods of organisation, management, optimisation, as well as efficient technical service (Koralewski & Wrona, 2014; Cisowski, 2016a, 2016b; Hao & Yue, 2016).

* Lublin University of Technology, Faculty of Mechanical Engineering, Institute of Technological Systems of Information, Nadbystrzycka Street 36, 20-618 Lublin,
e-mail: l.wojciechowski@pollub.pl
** Polish Air Force Academy, Department of Logistics, Dywizjonu 303 Street 35, 08-521 Dęblin,
e-mail: tadeuszc@poczta.onet.pl

A decision solution through operational research is a kind of procedure consisting of the following more important stages:

- examining the situation and the decision problem resulting from it,
- constructing the model describing a decision issue,
- searching a method of solution and solving the model of decision task,
- assessment of the correctness and feasibility of the obtained solutions, and the possible verification of the model and the method for solving,
- implementation of the solution and observation of the changes occurring under the influence of the said solution (Pawełczyk, 2005; Koralewski, Wrona & Wrona, 2016).

Considering the demand of the transport enterprises, it was deemed necessary to pick the topic of this article and to develop a new and efficient method consisting in the optimal selection of the operating parameters of the vehicles, which consequently leads to the increase of their durability in the utility process (Pires & Frazzon, 2016; Wojciechowski, Cisowski & Pazowski, 2015). It is known that the manner of the operation of a vehicle in any transport company significantly influences the efficiency of its economic activity, and the appropriate technical exploitation of transportation means aims at increasing their technical efficiency, which is necessary to perform their tasks (Rais, Alvelos & Carvalho, 2014; Koralewski & Wrona, 2015; Longwic, Lotko & Lotko, 2016).

2. SUBSYSTEM FOR DETERMINING THE EXTENT OF WEAR OF THE VEHICLES

Heuristics is considerably used by the subsystem for the algorithm used to rationalise the selection of the operating parameters of a motor vehicle.

The calculations start with defining the following sets:

- a set of transport tasks $T = \{T_1, T_2, \ldots, T_n\}$, where $n \in N$, where for every task T_i is defined:
  - T_{ipsxy} – starting point (cargo receipt point), where *x* and *y* are longitude and latitude respectively,
  - T_{ipdxy} – end point (cargo delivery point), where *x* and *y* are longitude and latitude respectively,
  - T_{il} (kg) – cargo to be transported in a given task (given in kg),
- a set of vehicles $V = \{V_1, V_2, \ldots, V_n\}$, where $n \in N$, where for every vehicle V_i is defined:
  - V_{ipxy} – vehicle's current location, where x and y are longitude and latitude respectively,
  - V_{ic} (kg) – vehicle's carrying capacity (given in kilogrammes),

- V_{iv} (*km/h*) – vehicle's technical speed (maximum speed, given in km/h),
- V_{is} (*1|0*) – vehicle's status (*value 1 – available, or 0 – unavailable*),

- a set of considered consumable parts $E = \{E_1, E_2, \dots, E_n\}$, where for every consumable part E_i are defined:
  - E_{iukm} (*%*) – percentage of wear and tear of parts per each kilometre driven,
  - E_{iuh} (*%*) – percentage of wear and tear of parts per each hour driven.

On their basis the following coefficients are determined:

- coefficient of wear and tear of parts per each kilometre driven:

$$E_{ukm} = \frac{\sum_{i=1}^{n}(E_{ukm})}{n}, \quad (1)$$

- coefficient of wear and tear of parts per each hour driven:

$$E_{uh} = \frac{\sum_{i=1}^{n}(E_{uh})}{n}. \quad (2)$$

For the aforesaid sets, the solutions are searched, which are specific allocation $f(T) = V$ (i.e. one car is assigned to perform every transport task, which is to be performed by this car) and:

- $V_{iT} = \{T_a, \dots, T_z\}$ is a set of tasks assigned to a vehicle,
- $V_{iR} = \{P_1, \dots, P_n\}$ is an ordered set of starting points and end points of all the tasks belonging to V_{iT},
- V_{iRt} is a cumulative journey time between the next points: $\{V_{ipxy} \rightarrow P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_n\}$,
- V_{iRd} is a cumulative distance between the next points: $\{V_{ipxy} \rightarrow P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_n\}$,
- Vehicle's initial load: $V_{iLP0} = 0$,
- Vehicle's load in point j is: $V_{iLPj} = V_{iLpJ} - 1 + P_{jT1}$, if P_j is a starting point of the task T_k,

or:

- $V_{iLPj} = V_{iLpJ} - 1 - P_{jT1}$, if P_j is an end point of the task T_k, where:
  - P_{jT1} is the cargo transported in the task T_k, whose starting point or whose end point is P_j, so $P_{jT1} = P_{Tk1}$,
  - In every P_j point the following relation is satisfied: $V_{iLPj} \leq V_{ic}$,
  - Cost value $C = V_{iRt} \times E_{uh} + V_{iRd} \times E_{ukm}$ is as low as possible,
  - Total cost value $C_t = n \sum i \{V_{iRt} \times E_{uh} + V_{iRd} \times E_{ukm}\}$ is as low as possible.

The next step is the optimal task assignment of vehicles and the optimal selection of the order of driving of particular vehicles between the starting points and end points of the tasks assigned to them. This problem is not NP-complete with at least exponential complexity, as it has much higher complexity than the problem of a travelling salesman that is only NP-complete. Therefore, in this method the heuristic algorithm was applied, which guarantees the finding of the solution in polynomial time and which produces a certain probability of finding the solution which is close to the optimum one or which is optimum in polynomial time.

3. SPATIAL DATA SERVER

Geoserver which was used in the algorithm is an open source server which enables geospatial data processing. It may publish data coming from all the leading spatial data sources using open standards. In practice, Geoserver returns the pictures in *.png format with the route alignment marked on them. The pictures generated in such a manner are added to the road map (OpenStreetMap). Thus, the process of generating the routes continues in the following manner:

- calculating the most favourable routes for particular cars on the basis of the data from the routing base;
- for every motor vehicle the further pairs of the route points are sent to Geoserver, with the demand to draw the route with the specified parameters between the said pair of points;
- in response, the next parts of the drawn routes are received in *.png format, which are subsequently marked on OpenStreetMap, as its new layers.

Developing new transport routes with Geoserver involvement was shown in figure 1.

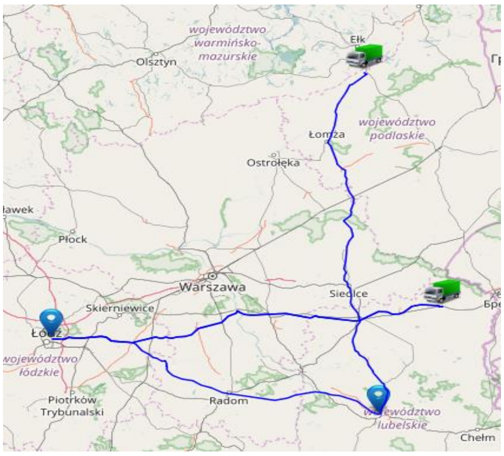


Fig. 1. Developing routes with Geoserver

In figure 2 the possibilities of the combination of the aforesaid tools are shown. It is a zoom of the crossroads of the routes of the vehicles, which perform their transport tasks, from the previous picture. It should be pointed out that the algorithm chooses the optimum routes very precisely and in detail on the basis of their real lengths and maximum permitted speeds.

Fig. 2. Indication of the details of Geoserver map

The level of the data detail of the routes marked on the map is so high that the numbers of the buildings, more important locations in the city, or the numbers of the route on which the vehicle is driving are visible (Fig. 3). It makes it possible to precisely locate the sending-receiving points and the position of the transport vehicles. During the application of the logarithm by a transport company, it is possible to analyse the routes chosen by the program.

Fig. 3. Location determination

4. DESCRIPTION OF THE COMPUTER INTERFACES TO ENTER THE DATA AND TO VISUALISE THE RESULTS

On the basis of the developed method of the optimisation of the performance parameters of the vehicle in the aspect of its durability, its simulator was performed in the form of a computer program – "Logistics" application. The program has an interface to enter the data and to visualise the results.

Entering the data includes the actions to specify:

- the data,
- the input parameters,
- the relations between the data,
- the goal function,
- the tasks to be performed,
- the logical objects.

The calculations include solving the given optimisation tasks, depending upon the expected result and the specific criteria and limitations. The presentation of the results is connected with the visualisation of the solutions of the calculations made. After starting the application, the main screen appears with the recently entered tasks of the means of transport operation (Fig. 4).

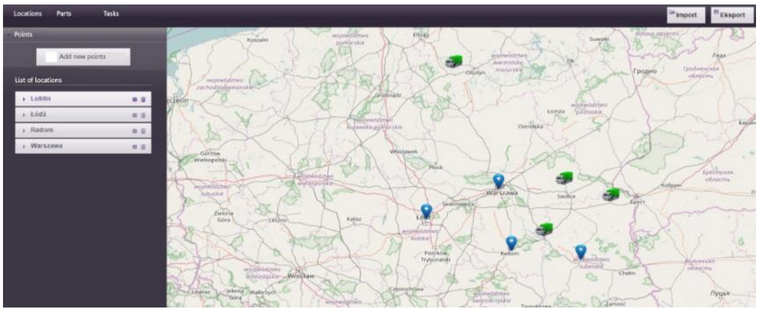


Fig. 4. Visualisation of "Logistics" application

On the left-hand side of the application (see Fig. 4) there is a user interface. It contains the respective tabs opening the right windows to enter the data, to specify their parameters, and to implement the calculation tasks. In the tab: "localisations" there are the fields to enter and define the sending-receiving transportation routes and the vehicles. To define the transportation routes their location should be marked on the map or their geographical coordinates should be entered. The identical activities should be performed with all the vehicles. All the added points and routes shall be visible on the map as well as in the list of the objects.

In order to add the means of transport, first, all the data characterising it should be entered (Fig. 5).

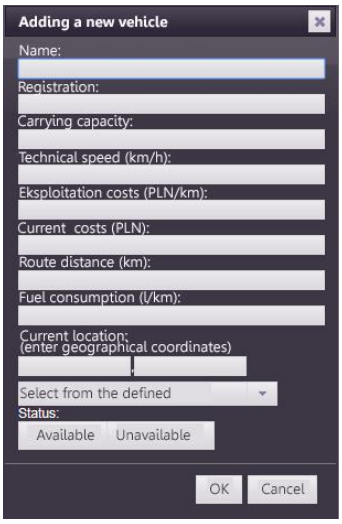


Fig. 5. Interface of adding a new vehicle

In the next tab "parts" (Fig. 6) the wear and tear of the vehicles' components in %/km or %/h may be added, i.e. the unlimited list of the parts which define their performance parameters. Depending upon the selected option – whether it is defining the route which is optimum for the transport of the cargo or analysing the exploitation of the vehicle's components, the algorithm finds the most favourable route which meets the imposed boundary conditions. When analysing the wear and tear of the parts of the vehicles, the algorithm examines if the extent of their deterioration was defined the most favourably in %/km or %/h. If their wear and tear is higher in %/km, the application searches for the shortest routes; however, if their wear and tear is higher in %/h – the algorithm chooses the fastest routes.

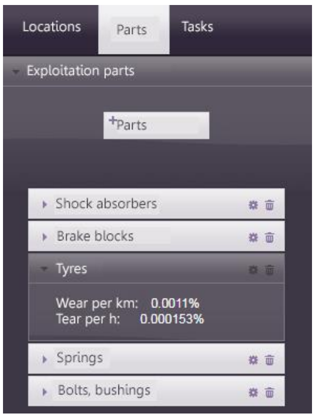


Fig. 6. Interface of adding an exploitation part

Planning the transport routes is placed in the field: "tasks" (Fig. 7).

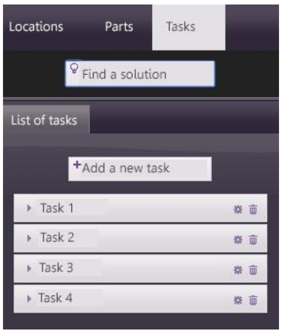


Fig. 7. A set of transport tasks

It is a set of all the transport tasks to be performed by the available vehicles. It defines the volume of the cargo which is to be transported by the vehicles, as well as the places of loading and unloading.

Additionally, the waiting time is added to the journey time and shown as the totalised value or as particular information: the journey time of "x", including the driving time of "y", and the waiting time of "z." In this case it is about the statutory working time of drivers. The algorithm considers it in the calculations and shows where the driver will be located in a given time, which enables to take a decision if they should be given another transport task. Figure 8 presents the visualisation of the calculations.

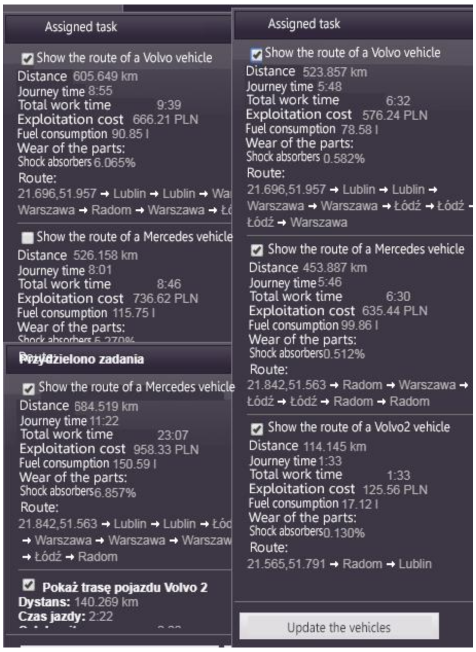


Fig. 8. Presentation of the results of the calculations of the algorithm

The above-mentioned example of the comparison of the final results of the calculations presents the possibility of showing different options of the results of the calculations, which enables any comparison thereof. It is essential for depicting the results of the calculations concerning the shortest, the fastest and the most optimal routes in respect of the defined consumable parts.

The tables of calculations presented in figure 8 are simultaneously visualised on the map (Fig. 9), on which the sending-receiving points, the transport vehicles, and different combinations of the routes proposed by the algorithm are placed. Depending upon the criterion taken, the routes are defined in such a manner to optimise the performance parameters of the vehicles.

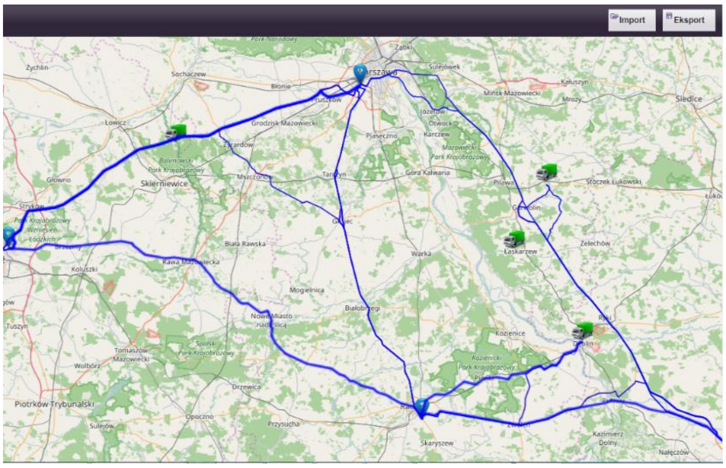


Fig. 9. Visualisation of the generated transport routes

5. VERIFICATION OF THE EXPERIMENTAL RESULTS AND THE REAL RESULTS ON THE EXAMPLE OF A SELECTED TRANSPORT ENTERPRISE

The real operation data on the vehicles from the fleet of the transport company that transports construction materials of homogeneous structure were used.

Table 1 below presents the list of the exploitation costs for three various options: the criterion of time, the criterion of road, and the real data from the company. The following concepts were compared: the route distance in relation to the cargo transported, the loaded distance of the goods, the fuel consumption, and the total exploitation costs. (Route Length Alg. – the calculations of the algorithm with the selection parameter of longer route but faster, Route Shortness Alg. – the calculations of the algorithm with the selection parameter of shorter route but slower, the Company – the real data from the transport activity of the transport company).

Tab. 1. List of the exploitation costs for different options

Option:	Route distance [kilometres]	Cargo transported [tonnes] (Q)	Fuel consumption [litres]	Exploitation cost [PLN]	Loaded distance of the goods [kilometres]
Route Length Alg.	411365.4	34428.59	85915.17	555343.35	27564.31
Route Shortness Alg.	374032.6	30380.05	78045.99	504943.96	25071.81
Company	457572	25803.66	95814.68	617722.22	30619.47

Figure 10 presents the exploitation costs of the vehicles depending upon their selected performance parameters. It should be pointed out that the reasonable selection of the performance parameters will bring the lower exploitation costs, the longer and more efficient work of the vehicle in the same time period.

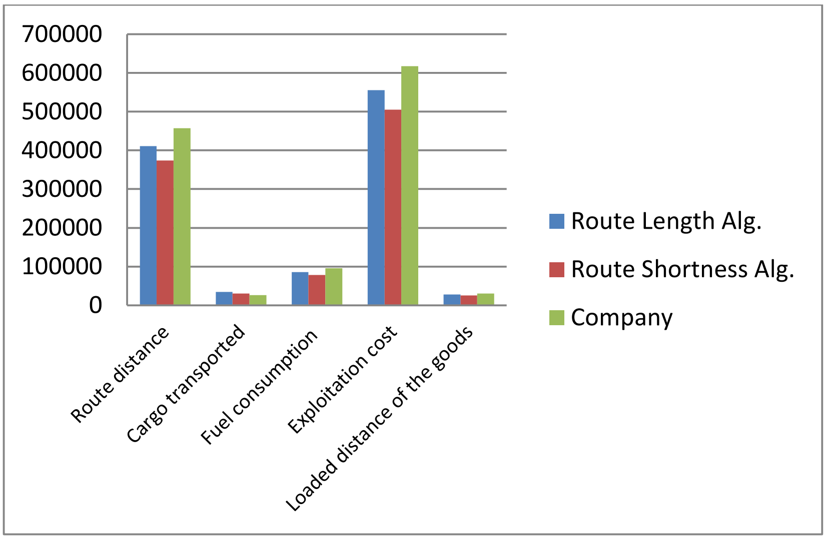


Fig. 10. The influence of the exploitation costs on the transport parameters

The next list shows the percentage wear and tear of selected parts during the exploitation of the motor vehicles. For this purpose, the aforesaid three options of exploitation were compared. As all the cases performed the same transport tasks, with the same number of the work readiness days and the same number of available means of transport, with the same load capacity, the total route distance made by the vehicles for each task could be averaged. For particular options of the calculations the distance of the route was as follows:

– for the real data gained from the transport company – 45757.2 km,
– in the case of the simulation with the criterion of the shortest routes selection – 37403.27 km,
– in the case of the simulation with the criterion of the faster but longer routes selection – 41136.53 km.

As it may be noticed, for all the regular service actions, the highest percentage wear and tear of the indicated parts was received when the transport company did not apply the tools for the optimal selection of the vehicles' performance parameters and the transport task planning. In the said case the wear and tear of the parts for the periodic vehicle inspection conducted every 20 thousand km was 228.79%, which meant around 2/3 time which was left to the third inspection. In the case of the periodic vehicle inspection every 60 thousand km 22.21% was left to the next inspection; in the case of the periodic vehicle inspection every 90 thousand km – 49.16%, and for the replacement of the parts every 150 thousand km – 69.34%.

The simulation of driving with the criterion of short routes showed much lower wear and tear of the parts, and it was respectively: for the replacement period every 20 thousand km – 187.02% (the second inspection is coming), every 60 thousand km – 63.59%, every 90 thousand km – 41.56%, and every 150 thousand km – 25.06%.

The last case of the calculations of the wear and tear of the parts of the vehicles concerned the simulation of driving with the criterion of faster but longer routes. The results of this simulation are as follows: for the replacement period every 20 thousand km – the wear and tear of the parts was 205.68%, for the replacement period every 60 thousand km – 69.93%, for the replacement period every 90 thousand km – 45.70%; whereas for the replacement period every 150 thousand km – 27.56%.

The aforesaid results of the research on the wear and tear of the vehicle's parts influencing its durability were depicted in figure 11.

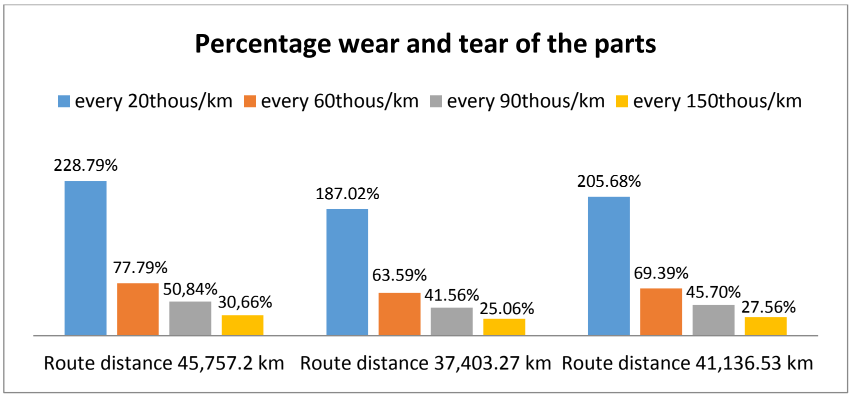


Fig. 11. Percentage wear and tear of selected consumable parts in three options of managing the car fleet

The results of the simulation research, in which the average route distance was 37403.27 km and 41136.53 km, are lower from the route distance of 45757.2 km, gained by the company which did not apply the tools for managing the transport process and the technical service of the vehicles, with the consideration of the appropriate selection of the performance parameters. The research showed that wearing the parts in shorter time results in choosing faster routes, usually longer in distance, the advantage of which is the optimum time of the performance of transport tasks and the bigger volume of the cargo transported. A faster route is usually identified with a shorter route.

6. CONCLUSIONS

This research aimed at presenting the method of optimising the performance parameters of a vehicle in the aspect of its durability, directed to the enterprises dealing with the transport of homogeneous goods. The research which has been conducted proves that optimising the vehicle's performance parameters to prolong the time of its reasonable exploitation is fully justified when using the developed algorithm. It was necessary to know the optimising tasks to be performed and the accepted limitation criteria in order to develop the algorithm correctly.

Therefore, the optimum selection of the performance parameters was found against:

- the delivery time,
- the efficient utilization of the load capacity of the vehicle,
- the reasonable exploitation of a transport unit.

The record of the limitations of the method was made regarding the functions of the criteria optimising the vehicles' performance parameters.

The obtained results allowed to form a reasonable method for optimising the performance parameters of the fleet in order to increase their reliability and efficiency during the performance of the transport tasks.

The practical part showed the selection of the optimum performance parameters which guaranteed the higher reliability of the vehicles in order to obtain the aforesaid assumptions.

REFERENCES

Rais, A., Alvelos, F., & Carvalho, M. S. (2014). New mixed integer-programming model for the pickup-and-delivery problem with transshipment. *European Journal of Operational Research*, *235*, 34–46.

Cisowski, T. (2016). Rating urban bus operation based on parameter complexity of their routes. *Autobusy: technika, eksploatacja, systemy transportowe, 17*(6), 821–824.

Cisowski, T. (2016). Rationalization of the carriage / collection of goods with one airport taking into account the time factor. *Autobusy: technika, eksploatacja, systemy transportowe, 17*(6), 1287–1293.

Hao, C., & Yue, Y. (2016). Optimization on Combination of Transport Routes and Modes on Dynamic Programming for a Container Multimodal Transport System. *Procedia Engineering, 137*, 382–390. doi:10.1016/j.proeng.2016.01.272

Koralewski, G., Wrona, J., & Wrona, R. (2016). Research of the city bus in his maintenance conditions. *Autobusy: technika, eksploatacja, systemy transportowe, 17*(11), 82–84.

Koralewski, G., & Wrona, R. (2014). Modelowanie ruchu autobusu miejskiego dla celów optymalizacji układu napędowego. *Logistyka, 3*, 3065–3074.

Koralewski, G., & Wrona, R. (2015). Ocena wpływu czynników eksploatacyjnych na algorytmy przełączania biegów autobusów miejskich. *Logistyka, 4*, 4151–4160.

Lotko, A., Lotko, M., & Longwic, R., (2016). Different multidimensional exploratory techniques in classifying variables into qualitative criteria of spare parts selection for passenger cars. *Advances in Science and Technology Research Journal, 10*(31), 185–193. doi:10.12913/22998624/64013

Pires, M. C., & Frazzon, E. M. (2016). On the research of linear programming solving methods for non-hierarchical spare parts supply chain planning. *IFAC-PapersOnLine, 49*(30), 198–203. doi:10.1016/j.ifacol.2016.11.167

Pawełczyk, M. (2005). Advisor – package for simulation tests of autonomous vehicles. *Prace Naukowe, seria Transport, 1*(21). Radom: Wydawnictwo Politechniki Radomskiej.

Wojciechowski, Ł., Cisowski, T., & Pazowski, P. (2015). Synthesis of decision models and algorithms in manufacturing systems. *Applied Mechanics and Materials, 791,* 108–115. doi:10.4028/www.scientific.net/AMM.791.108

Applied Computer Science, vol. 14, no. 4, pp. 25–33
doi:10.23743/acs-2018-27

Submitted: 2018-07-31
Revised: 2018-10-01
Accepted: 2018-12-01

extrusion, FEM simulation, compressed die

Irena NOWOTYŃSKA [0000-0001-5070-7619]*,
Stanisław KUT [0000-0003-0870-7548]**

COMPARATIVE ANALYSIS OF THE IMPACT OF DIE DESIGN ON ITS LOAD AND DISTRIBUTION OF STRESS DURING EXTRUSION

Abstract

In the paper a numerical comparative analysis of stresses in a steel die compressed by a ring during the extrusion process was presented. In the research, three design solutions of the die were used. The solutions vary depending on the quotient of the wall thickness of the die insert and the wall thickness of the compression ring while maintaining a constant tool diameter. The stresses occurring in the areas of the ring and the die were calculated depending on the design version of the tool and the pressure value. The analysis was carried out for the quotients of the die wall thickness to the ring wall thickness of 0.57, 1 and 1.75 and three press-in values of 0.004, 0.008 and 0.016 mm. The conducted research allowed determining the impact of the die design and assembly interference on the load bearing capacity. It was discovered that the use of a die insert with a smaller thickness compared to the thickness of the compression ring was the most advantageous from the point of view of the circumferential stresses.

* Rzeszow University of Technology, The Faculty of Management, Department of Computer Engineering in Management, Al. Powstańców Warszawy 8, 35-959 Rzeszów, Poland, e-mail: i_nowot@prz.edu.pl
** Rzeszow University of Technology, The Faculty of Mechanical Engineering and Aeronautics, Department of Materials Forming and Processing, Al. Powstańców Warszawy 8, 35-959 Rzeszów, Poland, e-mail: stan_kut@prz.edu.pl

1. INTRODUCTION

An application of an extrusion process requires the use of precise calculation methods to obtain specific technological objectives. Extrusion is characterized by high deformation values, large unit pressures on the tool surface, often complex geometry of the product. In this technology, in addition to giving the right shape of the product, it is necessary to take into account the economically justified life of the tool. The tool with such large unit pressures is exposed to wear on the work surface, elastic deformation and cracking. The life of the dies depends on their proper design and execution, taking into account the appropriate thermal and chemical treatment and the mechanical response to the material selected, the sequence of individual heat treatment operations, and the conditions under which the extrusion itself runs (tool temperature, insert geometry, process speed, type and amount of lubricant providing optimal tribological conditions). Another important parameter is the type of tool material used.

Because of the high unit pressures, the effort of the die material is often so high that it leads to premature wear and even destruction. Information on a tool life is particularly important when dealing with issues related to increasing the efficiency of the extrusion process. An important parameter for optimizing tool life is the introduction of a pre-compression stress in the die insert through the use of the compression ring (Groenbaek & Lund, 2008; Frater, 1989; Joun, Lee & Park, 2002; Yeo, Choi & Hur, 2001; Zimpel, 1996; Nowotyńska & Kut, 2014). Dies pre-compressed (reinforced) by a single ring are used mainly for extruding wire rods. Dies inserts can be made of steel or sintered carbide depending on the required abrasion resistance and pressure.

The aim of the article is to present the results of the comparative analysis of the impact of die design through the use of different thickness of rings and the size of the compression on the load carrying capacity of the tool in the extrusion process.

2. NUMERICAL MODELING OF THE EXTRUSION PROCESS

Numerical calculations were made using commercial MARC / Mentat from MSC Software. The extrusion process was modeled for a die compressed with one steel ring using three different quotients of the die wall thickness to the ring wall thickness (g_m/g_p = 0.57; 1 and 1.75) and three different press-in pressures (δ = 0.004 mm, 0.008 mm) and 0.016 mm). This allowed determining the most loaded places and values of stresses in individual areas depending on the design of the die and the size of the used interference. To determine the assembly insert δ, the following formula was applied:

$$\delta = \frac{D_{zw} - D_{wp}}{2} \tag{1}$$

where: D_{zw} – outer diameter of the die insert [mm],
D_{wp} – inner ring diameter [mm].

The die pattern with dimensions is shown in Figure 1. The angle of the work cone is assumed as $\alpha = 30^o$, while the length of the bearing $l_k = 5$mm.

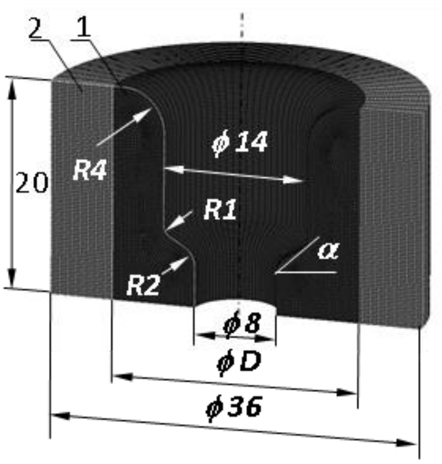


Fig. 1. Scheme of the die with dimensions (1 – die insert, 2 – compression ring)

A two-dimensional geometrical model of the process was constructed, which was analyzed with the assumption of axial symmetry (axial symmetrical model). An exemplary 3D expandable model with contact conditions for deformable bodies is shown in Figure 2.

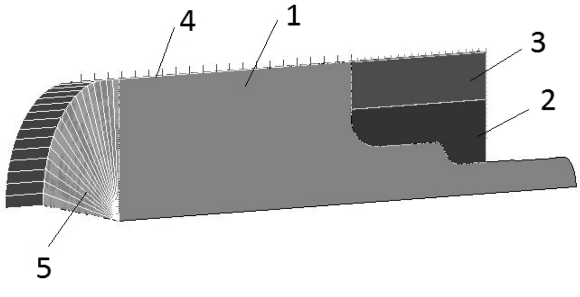


Fig. 2. Exemplary geometrical model of the analyzed extrusion process: 1 – formed material, 2 – die insert, 3 – prestressing ring, 4 – container surface, 5 – punch surface

The mechanical properties of the die (made of NC10 steel) are described by taking E = 210000 MPa and ν = 0.3. However, the properties of the extruded material are described by adopting an elastic-plastic body model with non-linear reinforcement. The values of the determined material constants for the extruded material were taken from the literature (Pater, 2003) and are given in Table 1. The friction model is described by Coulomb's law. The coefficients of friction between the extruded material and the rigid surfaces of the punch and the container and the deformable die were assumed to be equal of 0.3. On the other hand, the coefficient of friction between the die insert and the compression ring was 0.1 (Kut & Nowotyńska, 2011).

Tab. 1. Mechanical properties of the tested material

Material	Strain intensity ε	Yield stress σ [MPa]	Young's modulus E [MPa]	Poisson's coefficient ν
Lead (99.98%)	0	5	18 000	0.42
	0.05	10.42		
	0.10	13.27		
	0.15	15.59		
	0.20	17.30		
	0.30	18.78		
	0.40	18.72		
	0.50	18.64		
	1	19.05		

For the construction of finned elements of deformed material, the elements of the class 10 type – axisymmetric ring quadrilateral were used. Numerical simulation was performed using the global remeshing option. The size of the inserts and the compression ring was about 0.25 mm while the extruded material was about 0.4 mm.

In the paper an analysis of the circumferential stresses depending on the g_m/g_p quotient for three different assembly pressures was done. The circumferential stresses prevailing in the die insert and the ring during extrusion were determined.

3. CALCULATION RESULTS AND THEIR ANALYSIS

Often, in industrial practice, conical dies are used in the extrusion process. Despite many advantages of using such a die geometry, there is a danger of wear and tear even in the input area and at the calibration cone. This is due, inter alia, to the occurrence of such circumferential stresses in such dies. One way to reduce the damaging effects of these stresses on tools during extrusion is to use pre-compressed dies.

The purpose of the numerical tests was to determine the impact of the compressed die construction, expressed in a different ratio of the die insert thickness (g_m) to the ring thickness (g_p) and the size of the interference per peripheral stress value in the extrusion process. The compressed die tests were carried out for three cases of interference (1) which, as already mentioned, were: 0.004 mm; 0.008mm; 0.016 mm and three different ratios of die thickness to the thickness of the spring ring, i.e. g_m/g_p to 0.57; 1; 1.75. Stress readings were made at 7 points for all dies and inserts in the die insert and ring (Fig. 3).

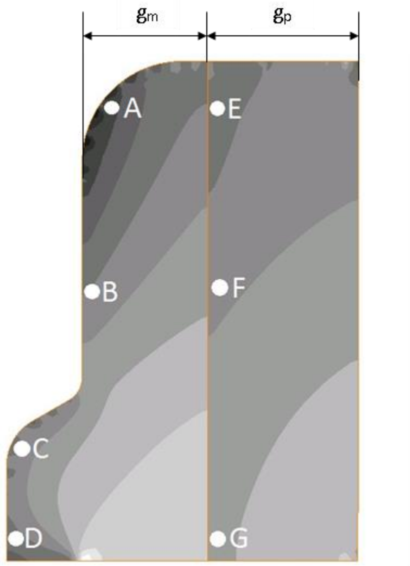


Fig. 3. Locations of measurement points of circumferential stresses

Figures 4 and 5 present the relationship of the circumferential stress prevailing in the die insert and the compression ring depending on the applied die design variant (g_m/g_p) for the three tested assembly pressures. By applying the appropriate die design realized by changing the thickness of the die insert and the spring ring, it is possible to influence on the values of maximum of peripheral stresses occurring in the tool during extrusion. From this point of view, the most beneficial is an application of the variant where $g_m/g_p = 0.57$ as the smallest values of circumferential stress in all the analyzed areas of the tool occurred, including negative stresses.

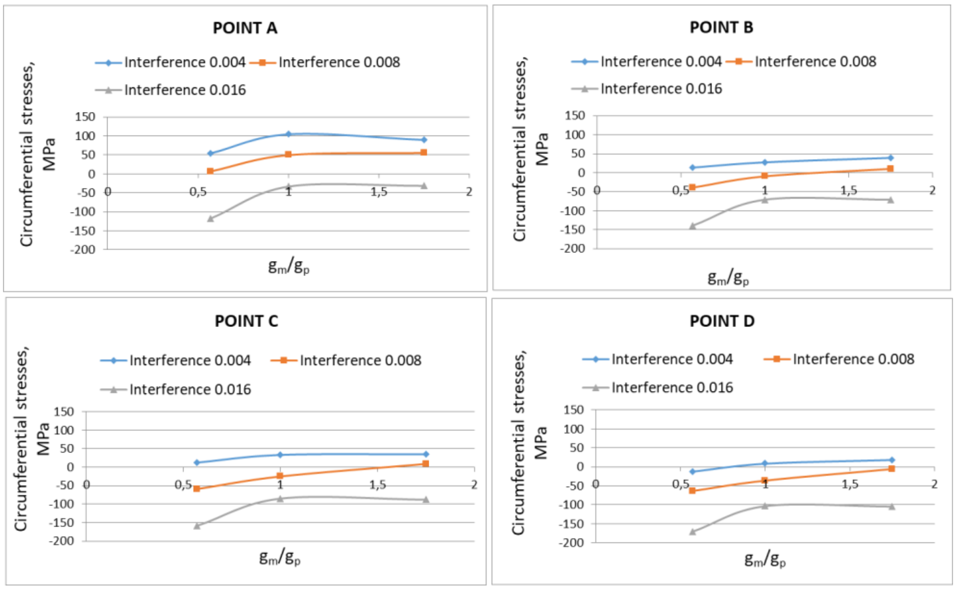


Fig. 4. The impact of the g_m/g_p quotient on the value of maximum circumferential stresses in individual areas of the die insert depending on the used assembly interference

With the use of pressure of 0.008 mm, the smallest positive values of circumferential stresses of 7 MPa and the highest negative values of these stresses equal to –170 MPa for the measuring point D with the pressure value equal to 0.016 mm were recorded in the area A. From the point of view of die durability C and D areas are significant as they are places most vulnerable to damage. In the case of the measuring area C, the study of the influence of different geometrical variants on circumferential stresses showed that at the pressure of 0.004mm and the geometrical ratio $g_m/g_p = 0.57$, the circumferential stresses reached the value of 13MPa. The use of the variant where $g_m/g_p = 1.75$ caused an increase of over two and a half times in the value of peripheral stresses and amounted to 35 MPa.

Figure 5 shows the calculated stress values in the area of the compression ring. In all considered geometrical cases and the areas of research, i.e. E, F, the values of peripheral stresses are positive. In particular measuring areas and the corresponding pressures, the stress values are similar, e.g. for the 0.004 mm pressure, they oscillate between 87–96 MPa for the area E. The differences occur only when different pressure values are used. The smallest values of circumferential stresses were noted for the 0.004 mm pressure when using a geometric variant where the die insert and the ring were of the same thickness. Hence, in the case of a compression ring, no significant change in the geometric variant was affected.

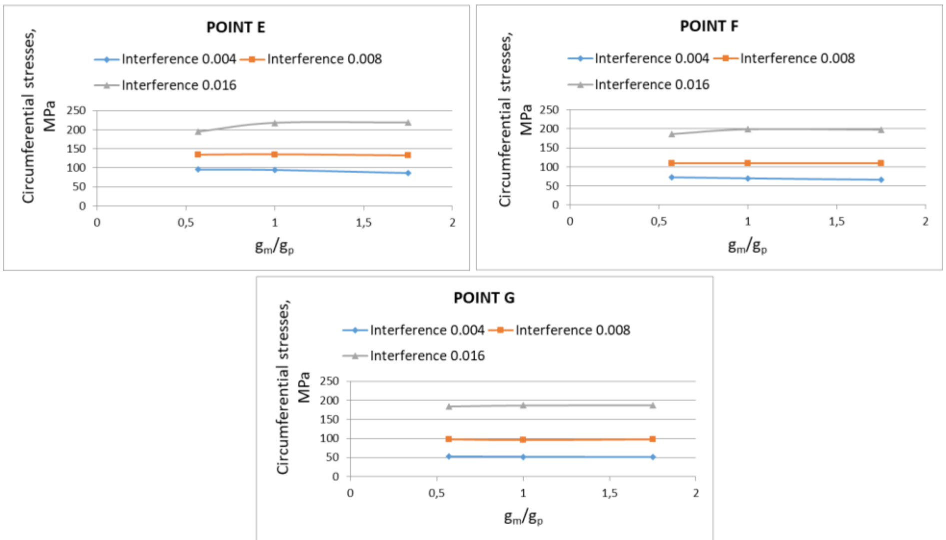


Fig. 5. The impact of the g_m/g_p quotient on the value of maximum circumferential stresses in individual areas of the compression ring depending on the used assembly interference

When comparing the values of circumferential stresses in the ring area with their adopted geometry, one can conclude that they are larger in the measuring areas of rings and range from 51 MPa and pressures 0.004 mm to 219 MPa for 0.016 mm for the variant $g_m/g_p = 1.75$.

When considering tool lifetime, apart from the values of peripheral stresses, the distributions and gradients are also important in the points selected for the analysis. Figure 6 presents examples of stress distribution results in a die compressed by a steel ring and the ring itself for three geometrical variants and exemplary pressures of 0.008mm and 0.016mm. The use of a pre-compressed die with one steel ring, depending on the different geometric variant, i.e. a different ratio between the thickness of the die insert and the compression ring (Fig. 6), changes the distribution of stresses mainly in the die material. For the variant where the thickness of the die is slightly smaller than the rings, one can notice the concentrated stresses at the die entrance as well as passing through the calibration strip. In contrast, there is no visible large stress gradient on the cross-section of the tool, i.e. a large variation in the magnitude of radial stresses. The die material in the whole area is more evenly loaded, which consequently at the same load causes a reduction in the level of stress values, which is very beneficial from the point of their overage and operational durability.

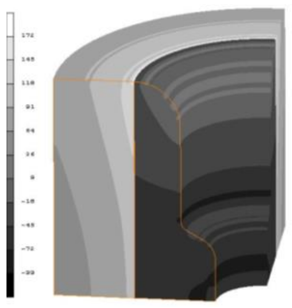
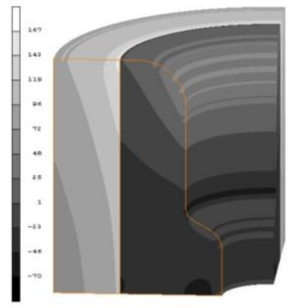
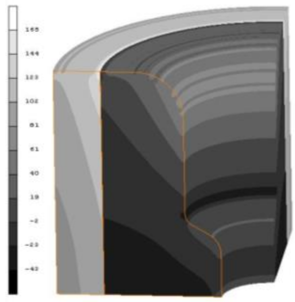
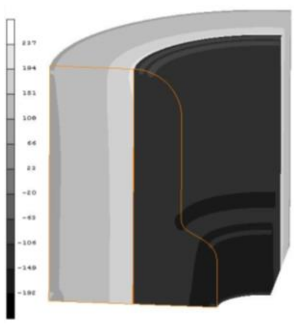
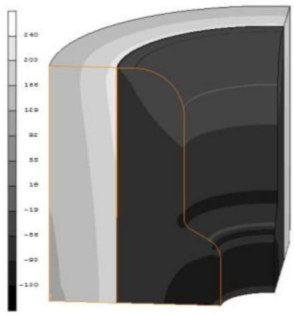
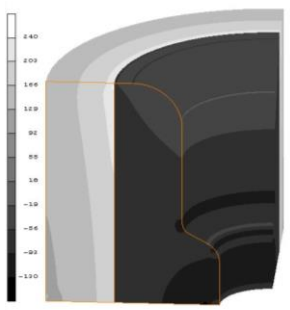
$g_m/g_p = 0.57$	$g_m/g_p = 1$	$g_m/g_p = 1.75$
Interference 0.008 mm		
		
Interference 0.016 mm		
		

Fig. 6. An example of distributions and values of peripheral stresses in the tool during extrusion with pressing 0.008 mm and 0.016 mm the use of g_m/g_p amounting to 0.57, 1 and 1.75

The use of die pre-compressed with a ratio of $g_m/g_p = 1.75$ causes the occurrence of stress differentiation on the cross section of the die material. Along with the change in the dimensional parameters of the die and the ring, the stress distribution also changes depending on the pressure value. This is particularly evident in the geometric variant $g_m/g_p = 0.57$. In this case, an increase of the pressure from 0.008 to 0.016 mm leads to a more even distribution of stresses in the entrance and exit areas of the die. In engineering practice, when selecting the interference in the case of the insert and compression ring made of steel, special attention should be paid to the fact that the stress strains between the insert and the ring occur as little as possible during the load. In addition, the values of circumferential tensile stresses in the ring should not exceed the permissible values. Therefore, it is important to look for such solutions, including geometric ones, in order to minimize unfavorable stress gradients.

4. CONCLUSIONS

Thanks to numerical modeling of real processes, in engineering practice one can choose the best variant of the technological process ensuring the longest possible tool life and high accuracy and repeatability of the dimensional and shape of products while limiting costly experiments. The numerical method applied allows the determination of stress values in a tool of any profile, taking into account differential stresses from extruded material and friction conditions. The obtained results showed the existence of a relationship between the geometrical parameters of the tools (expressed in the change in the thickness of the die insert and the compression ring) and the values of the assembly pressures and the values of peripheral stresses in pre-compressed dies.

REFERENCES

Groenbaek, J., & Lund, E. (2008). Tool Optimization by Means of Effective Prestressing System. New Delhi, India: Conference Asia Forge.

Frater, J. L. (1989). Application of Finite Element Methods to the Design of Prestressed Tooling. *Journal Mater. Shaping Technology*, *7*(1), 49–55. doi:10.1007/BF02834802

Joun, M. S., Lee, M. C., & Park, J. M. (2002). Finite element analysis of prestressed die set in cold forging. *International Journal of Machine Tools & Manufacture*, *42*, 1213–1222.

Kut, S., & Nowotyńska, I. (2011). Comparative numerical analysis of die wear during extrusion process of metals with different properties. *Hutnik, Wiadomości Hutnicze*, *78*(11), 925–929.

Nowotyńska, I., & Kut, S. (2014). Analiza numeryczna naprężeń podczas wyciskania w matrycy zwykłej i sprężonej. *Hutnik, Wiadomości Hutnicze, 81*(7), 465–469.

Pater, Z. (2003). Ołów jako materiał modelowy do symulacji procesów obróbki plastycznej na gorąco. *Obróbka Plastyczna Metali*, *14*(4), 41–48.

Yeo, H. T., Choi, Y., & Hur, K. D. (2001). Analysis and Design of the Prestressed Cold Extrusion Die. *The International Journal Advanced Manufacturing Technology*, *18*(1), 54–61.

Zimpel, J. (1996). Obliczanie matryc wzmacnianych jednym pierścieniem. *Obróbka Plastyczna Metali*, *2*, 43–56.

Applied Computer Science, vol. 14, no. 4, pp. 34–53
doi:10.23743/acs-2018-28

Submitted: 2018-09-30
Revised: 2018-12-07
Accepted: 2018-12-16

pressure vessel, finite element, simulation, numerical analysis, tank

Paweł BAŁON [0000-0003-3136-7908]*/***, *Edward REJMAN*[**], *Bartłomiej KIEŁBASA*[*], *Janusz SZOSTAK* [0000-0002-7789-3383]***, *Robert SMUSZ* [0000-0001-7369-1162]****

NUMERICAL AND EXPERIMENTAL ANALYSIS OF THE STRENGTH OF TANKS DEDICATED TO HOT UTILITY WATER

Abstract

The focus of this paper are experimental and numerical strength tests of domestic hot water storage tanks. The tests involved the verification of the minimum wall thickness for the assumed operating parameters while meeting all safety standards. The authors presented numerical and experimental analyses for the verification of strength parameters of axial cylindrical tanks due to the lack of methodological guidelines for this type of equipment. In order to verify the conducted theoretical considerations and calculations, experimental tests of samples of front welds produced with austenitic steel as well as a pressure test for the whole tank were conducted using a research test stand.

[*] SZEL-TECH Szeliga Grzegorz, Wojska Polskiego Street 3, 39-300 Mielec, Poland, e-mail: p.balon@szel-tech.pl, e-mail: bartek.kielbasa@gmail.com
[**] Rzeszów University of Technology, The Faculty of Mechanical Engineering and Aeronautics, Department of Mechanical Engineering, Powstańców Warszawy Avenue 9, 35-959 Rzeszów, Poland, e-mail: erejman@prz.edu.pl
[***] AGH University of Science and Technology, Faculty of Mechanical Engineering and Robotics, Department of Manufacturing Systems, Mickiewicza Avenue 30-B4, 30-059 Kraków, Poland, e-mail: szostak@imir.agh.edu.pl
[****] Rzeszów University of Technology, The Faculty of Mechanical Engineering and Aeronautics, Department of Thermodynamics, Powstańców Warszawy Avenue 9, 35-959 Rzeszów, Poland, e-mail:robsmusz@prz.edu.pl

1. INTRODUCTION

A pressure vessel is a reservoir manufactured to contain fluid (liquid or gas) at a pressure substantially different from the ambient. Cylindrical pressure vessels have widespread industrial applications and have become a type of equipment widely used in industry and everyday life. They are used in power plants, nuclear reactors, chemical processing reactors, and food industries. They appear in industrial compressed air receivers and domestic hot water storage tanks. Other application areas of pressure vessels are recompression chambers, distillation towers, autoclaves, oil refineries, petrochemical plants, vehicle airbrake reservoirs, and storage tanks for liquefied gases such as ammonia, propane, butane, and LPG, etc. They often perform under extreme pressure and temperature conditions (Li, Sheng & Zhang, 2012; Lakshmi Devi & Hari Shankar, 2016). Fundamental loads acting in the tanks are internal hydrostatic pressure and internal uniform pressure. The design of the pressure vessels is mainly related to their strength and the strength study mainly includes stress concentration analysis in the neighborhood of the head and cylindrical shell joint and in the heads of the tank. Pressure vessels can be closed at the ends by different shapes of the heads from flat plates to hemispherical domes. At the junction between a cylindrical shell and a vessel head there is discontinuity of meridional curvatures. These discontinuities of curvatures disturb the membrane stress state and have significant influence on the strength of the structure. This issue has been analyzed in books (Ziółko, 1986; Harvey, 2000) and standards (PN-EN 1993-1-6:2007, 2007; PN-EN 1993-4-2:2009, 2009). One of the most important topics in design is optimal shaping of the entire structure of the tanks. In many works the problem of the optimal tank shape was analyzed (Ventsel & Krauthammer, 2001; Błachut & Magnucki, 2008; Lewiński & Magnucki, 2010, 2012) and special attention was paid to the junction between the vessel and its head because this region is usually subject to considerable stress concentration due to the edge effect (Krużelecki & Proszowski, 2012).

The aim of this work was to describe the strength performance of a cylindrical vertical domestic hot water storage tank. Numerical calculations and experimental strength tests were made to verify the wall thickness for assumed work parameters while meeting all safety standards.

The correct assessment of strength properties requires both numerical and experimental testing of the tank, including its critical nodes, e.g. welded joints. In the first stage of strength tests, tensile tests of samples, tests of mechanical parameters of material and tests of welded joints were conducted. Then an analytical estimation of peripheral and radial stress values was carried out. In the second stage, numerical analysis of the model's real object mapping was performed using the MES software. In the third stage, in order to assess the strength of the structure, strain measurements were done in the shell of the tank subjected to pressure.

The main purpose of the experimental tests was to obtain the data to assess the strength of the tank subjected to the internal pressure load. Assessment of strength properties requires experimental tests of both: tank and critical nodes, e.g. welded joints. In the first stage of strength tests, tensile tests of welded butt joints were carried out. In the second stage the structural behavior of a cylindrical tank subjected to internal pressure was investigated.

2. TESTS OF MECHANICAL PROPERTIES OF SHEET USED ON VESSELS

2.1. Determination of the normal anisotropy coefficient R

A vessel is a device that consists of two main parts: the bottoms and the shell. While the shaping of the shell does not cause any technological problems, forming the bottoms is a serious problem. This is particularly important when shaping products with large diameters requiring expensive and heavy dies (Bałon & Świątoniowski, 2016a, 2016b; Bałon, Świątoniowski, Szostak & Kiełbasa, 2016). For this reason, the properties of the sheet material must be very precisely defined for the description of the stamping process. One of the most important parameters is the normal anisotropy coefficient R. Therefore, research on this parameter has been carried out.

The normal anisotropy R ratio by Lankford is defined as the ratio of transverse strain increments during uniaxial stretching. Assigning the main directions for flat stress state of indices: 1 – direction of rolling, 2 – direction perpendicular to the rolling direction in the plate plane, 3 – normal direction to the plate surface and determied by de_{ij} the tensor components of the plastic strain increment, this coefficient – for the sample cut at an angle α to the direction of rolling – which we can express in the form of a quotient (Bałon & Świątoniowski, 2016a, 2016b; Bałon & Świątoniowski, 2013).

$$R_\alpha = \frac{de_{22}}{de_{33}} \tag{1}$$

and respectively:

$$R_0 = \frac{de_{22}}{de_{22}} \text{ and } R_{90} = \frac{de_{11}}{de_{33}} \tag{2}$$

In practice, with homogeneous and proportional deformations, in determining the anisotropy coefficients, in the place of de_{ij} increments, the final values of real strains can be used, so the final formula will take the form (Dyrektywa 97/23/WE, 1997; Rozporządzenie Rady Ministrów, 2002):

$$R_\alpha = \frac{ln\frac{b_1}{b_0}}{ln\frac{g_1}{g_0}} \quad (3)$$

where: b_0, b_1 – widths of the sample before and after deformation,
g_0, g_1 – thicknesses of the sample before and after deformation.

The measurement results allow us to state that the value of the steel yield point determined at 100 °C with a disproportionate elongation $R_{p0,2}$ [MPa] varies from 276.1 MPa (for the 0° direction samples) to 283.5 MPa (for the 90° direction samples), which indicates no significant influence of the directivity of the material structure on its properties after rolling. This is also confirmed by the results relating to the value of immediate tensile strength R_m – 548.8 MPa (for the 0° direction samples) and 551.3 MPa (for the 90° direction samples). In turn, the damage elongation measured after assembling both parts of the sample reaches 53%.

Measurements of the strains of the samples during thier uniaxial stretching were used to determine the normal anisotropy coefficient R. The mean R_{sr} value was calculated using the dependence:

$$R_{sr} = \frac{1}{4}(R_0 + 2R_{45} + R_{90}) \quad (4)$$

where: R_0, R_{45}, R_{90} – values of normal anisotropy coefficients for the direction according to the direction of rolling, directed at an angle of 45° and transverse to the direction of rolling

The average value of the normal anisotropy coefficient in a steel sheet of DIN 1.4541 is R_{st} = 0.9832. Therefore, the value of the anisotropy coefficient for the sample cut at an angle can be described by the formula:

$$R_\alpha = \frac{ln\frac{b_1}{b_0}}{ln\frac{l_0 g_0}{l_1 g_1}} \quad (5)$$

where: b_0, b_1 – measured lengths of the sample before and after elongation.

On the basis of the conducted tests, it was found that the tested steel sheet DIN 1.4541 does not show normal anisotropy to a degree that can significantly affect the course of the pressing process. In the most general case, during stamping of the bottom from a flat disc, in the flange zone, the material is subjected to radial tensile stresses and peripheral compressive stresses, and thus the product of main stresses fulfills the unevenness $s_{11}s_{22} < 0$.

In this case, according to the condition of metal transitioning into the plastic state of Mises-Hill:

$$s_{11}^2 - \frac{2R_0}{1+R_0} s_{11} s_{22} + \frac{R_0(1+R_{90})}{R_{90}(1+R_0)} s_{22}^2 = s_p(1) \tag{6}$$

the absolute decreases, and the necesssity for plasticizing stress values $|s_{11}|$ and $|s_{22}|$ could be expected only at $R > 1$.

The bottoming operation is a cold process, hence the deformation of the material formed in subsequent phases leads – through dislocations in the crystal lattice – to the strengthening of the blank material. This phenomenon must be taken into account when determining the process parameters.

In the macroscopic analysis, the non-linear relation between stress σ, and strain φ describes the metal strengthening curve, the form of which depends on the adopted model of strengthening – isotropic or kinematic (Bałon, Świątoniowski, Szostak & Kiełbasa, 2016, 2017; Bałon, Świątoniowski & Kiełbasa, 2017).

2.2. Isotropic hardening model

In the stamping process, the kinematic model better reflects the physical aspect of the proces than the isotropic model. In the considered case of stamping, in which the load increases continuously up to the maximum value, there is no need to take into account stress hysteresis and therefore, the sufficient accuracy is ensured by a much simpler isotropic model.

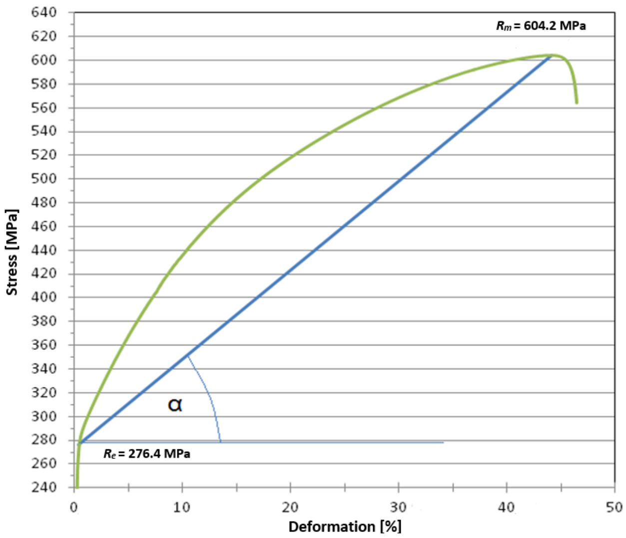


Fig. 1 Hardening curve of DIN 1.4541 steel with applied connecting section $R_{p0,2}$ and R_m (blue color) – isotropic hardening

The nonlinear material with isotropic hardening is defined by the value $R_{p0,2}$ and the value R_m, while the section between $R_{p0,2}$ and R_m is defined by the tangent of the angle of inclination of the α curve. Transformation of the stiffness matrix takes place once strain or stresse values do not exceed the value specified by $R_{p0,2}$ and above this value each time at successive iterations.

Hardening curve of the material determined with extrapolation according to the method of Krupkowsky-Gesetz:

$$\sigma_w = K * (\varphi + \varphi_0)^n \tag{7}$$

where: n = 0.31500, φ_0 = 0.00055, K = 2442.194.

3. EXPERIMENTAL RESEARCH OF FRONT WELDS

Steel 1.4541 (EN designation X6CrNiTi18-10), which was used in fabrication of the tank belongs to the largest austenitic stainless steels group with high corrosion resistance, and can be welded in all dimensions without becoming susceptible to intergranular corrosion. The chemical composition of the steel 1.4541 shown in Table 1.

Tab. 1. Chemical composition of the X6CrNiTi18-10 in %

Element	**C**	**Si**	**Mn**	**P**	**S**	**Cr**	**Mo**	**Ni**	**Cu**	**Ti**
Standard	0.08	–	2	0.045	0.015	17–10	–	12	–	0.7
Tested	0.034	–	1.20	0.021	0.001	17.21	0.057	9.12	0.10	0.394

The analysis concerns the mechanical properties of the TIG welding butt joints with argon shielding. Test samples were prepared from 2.0 mm thick sheets of steel 1.4541.

In accordance with standards, a tensile test was carried out on a certified strength machine and the results of the strength tests performed on welded joints made of steel 1.4541 under quasi-static conditions are shown in Table 2.

Tab. 2. Mechanical properties of the 1.4541 steel obtained in the tensile tests

Test No.	**Ultimate tensile strength [MPa]**	**Average ultimate tensile strength [MPa]**	**Yield strength [MPa]**	**Average yield strength [MPa]**
1	615	612.2	261	262.2
2	618		258	
3	608		269	
4	604		271	
5	616		252	

The tests show that destruction and broken areas of all samples occurred outside the heat affected zone and the tensile strength of the joint was higher than the base material. Furthermore, ultimate tensile strength (UTS) of the steel 1.4541 samples was higher than normative values of the UTS (normative UTS of the steel 1.4541: 520 MPa).

In the next step, the experimental tests on an internally pressurized tank were carried out. The subject of the experimental test was a cylindrical pressure vessel with a mean diameter of the cylindrical part equal 480 mm, made of steel X6CrNiTi18-10. The wall thickness of the cylindrical part of the tank and the top bottom head walls is the same and equal 2 mm and the main geometric dimensions of the head geometry are shown in Fig.2. The heads of the tank were manufactured by spinning and the shape of the top and middle surface of the head was defined by spline curve. The bottoms were manufactured according to the DIN 28013 norm (Warunki Urzędu Dozoru Technicznego WUDT/UC/2003, 2005; Bałon, Świątoniowski, Szostak & Kiełbasa, 2016).

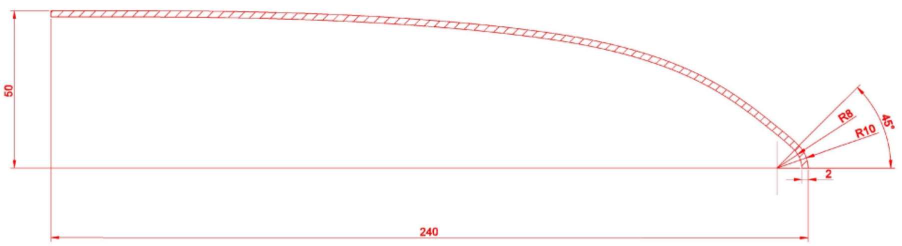


Fig. 2. Head tank geometry

For the joining of samples, the classic TIG method and the method with argon bluing on the side of the root of the weld were used. In the case of stretching samples made of X6CrNiTi18-1 steel made with the TIG method without argon bluing from the side of the ridge, the sample was destroyed in the cross-section of the weld, and for welding with the bluing – outside the weld area (Bałon & Świątoniowski, 2014).

Based on the above tests, it can be concluded that, with regard to the same welded material, the TIG welding technology using argon blown from the root side of the weld significantly improves the quality of the weld surface as well as its tensile strength. The average value of tensile strength of the connection was R_m = 630.6 MPa.

For this reason, the TIG method was proposed for the welding of the vessel, with argon bluing on the ridge side (Bałon, Świątoniowski & Szostak, 2015; Bałon & Świątoniowski, 2014).

Fig. 3. Welded sample after tensile test –X6CrNiTi18-1 (1.4541) steel – TIG welded sample with argon bluing on weld ridge

4. ANALYTICAL AND NUMERICAL STREGTH CALCULATION OF THE VESSEL

The designed tank is dedicated to storing water with a temperature of up to 100 °C and a working pressure P_1 = 6 bar equal to the pressure of water feeding from the network. Nominal, design and trial pressures were adopted in accordance with the regulatory literature in force in the European Union and Poland.

The basis for the design of the tanks is P_s pressure which is the maximum allowable pressure specified by the producer for which the device has been designed (Dyrektywa 97/23/WE, 1997; Rozporządzenie Ministra Gospodarki, 2005; PN-EN 13445-1, 2014; PN-EN 13445-3, 2014).

The following assumptions were made for the calculation of the vessel:

- V = 192 dcm^3 – volume of water in the vessel,
- P_1 = 6 bar – water pressure in the vessel corresponding to the supply pressure (working pressure),
- T_o = 10 °C = 277 °K – minimum water temperature,
- T = 70 ºC = 343 °K – maximum water temperature.

Taking into account fluctuations in the water pressure in the network and the tolerance of pressure regulation in safety valves, P_S = 7.5 bar was assumed:

Hydrostatic fluid pressure:

- H = 1.1 m – the height of the liquid column,
- ρ_w = 1000 kg/m^3 – water density.

Hydrostatic pressure:

$$P_h = H * \rho_w * g = 1.1 * 1000 * 9.81 = 1.079 * 10^4\ P \tag{8}$$

Due to the negligibly low hydrostatic pressure, which is approx. 1.8% of the working pressure, its influence on the design pressure was not taken into account.

Design pressure (PN-EN 13445-1, 2014):

Base on (PN-EN 10131:2006, 2006), the design pressure P_d should meet the condition:

$$P_d = 1.5\ P_S = 11.25\ bar \tag{9}$$

Test pressure:

The test pressure P_{test} should be:

$$P_{test} = 1.43\ P_S = 10.72\ bar \tag{10}$$

The higher pressure from P_d and P_{test}: should be taken as the design pressure:

$$P_d = 11.25\ bar \tag{11}$$

For the assumed bottom geometry and pressure of P_d = 11.25 bar, numerical calculations of the bottom using the COMSOL program were performed. The maximum reduced stresses reached the value of 210 MPa, which is lower than the material yield point of about 260 MPa, so the structure is safe.

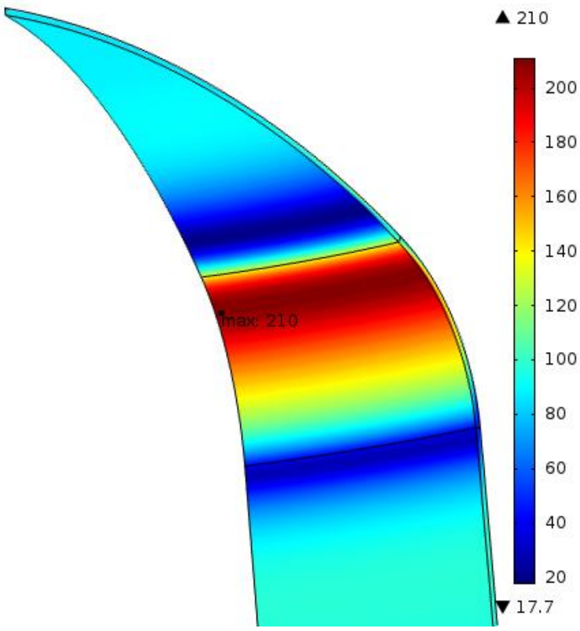


Fig. 4. A state of reduced stresses in the bottom according to DIN 28013 loaded with the design pressure of P_d = 11.25 bar, *e* = 2 mm – thickness of the wall (COSMOL)

In the case of vessel design according to (Rozporządzenie Ministra Gospodarki, Pracy i Polityki Społecznej, 2003), it is necessary to determine the design temperature T, which affects the strength properties of the vessel material. It is determined with dependencies:

$$T = T_C + 20°C \tag{12}$$

where: T_C – maximum temperature of the medium.

$$T = 70°C + 20°C = 90°C \tag{13}$$

According to (PN-EN 13445-1, 2014), the permissible stress f_d is the criterion stress for assessing the force of the vessel material. Using the design method consistent with the recognized engineering practice, and having the material from which the vessel will be made, in accordance with EN 10088-3: 2005, the strength properties of the tank material were adopted as follows:

$$R_{p1,0} = 260\ MPa \tag{14}$$

$$R_M = 540\ MPa \tag{15}$$

$$A_{min} = 35\% \text{ – elongation} \tag{16}$$

It should be stated that the material parameters obtained during the tests significantly exceed the values given by the standard. Due to the fact that the designed tank works at a temperature of T = 70 °C, it can be estimated on the basis of the EN 10088 norm that the appropriate mechanical properties for the tested steel are:

$$R_{p1,0/T} = 273\ MPa \tag{17}$$

$$R_{M/T} = 575\ MPa \tag{18}$$

The permissible stresses f_d are determined on the basis of chapter 6.4 according to (PN-EN 1993-1-6:2007, 2007).

$$f_d = \max[\left(\frac{\frac{R_{p1,0}}{T}}{1,5}\right); \min\left(\frac{\frac{R_{p1,0}}{T}}{1,2}\right); \left(\frac{\frac{R_m}{T}}{3}\right) \tag{19}$$

$$f_d = \max[(145{,}3); \min(181{,}7); (165)] \quad (20)$$

$$f_d = 165\ MPa \quad (21)$$

The *f* stresses in the structure should satisfy the following conditio:

$$f \leq f_d \quad (22)$$

4.1. Calculation of cylinder wall thickness of a tank part according to peripheral stresses

$$e = \frac{P*D_i}{2*f*z-P} \quad (23)$$

or

$$e = \frac{P*D_e}{2*f*z+P} \quad (24)$$

where: $P = P_d = 11.25$ MPa – design pressure,
e – required wall thickness,
z – connector factor (Dyrektywa 97/23/WE, 1997):
z = 1 – when the tested object together with the weld is subjected to destructive and non-destructive tests,
z = 0.85 – when the object with the weld is subjected to random non-destructive tests,
z = 0.7 – for non-destructive visual tests.

Considering that the assessment of weld properties was done by testing a welded metal sample and ther are conditions for visual assessment of the weld in the structure, *z* = 0.85 was accepted for the calculations.

For the internal diameter of the tank of D_e = 480 mm, the theoretical wall thickness of the tank was determined as follows:

$$e = \frac{1{,}125\frac{N}{mm^2}*480mm}{2*165\frac{N}{mm^2}*0{,}85=1{,}125\frac{N}{mm^2}} = 1.920\ mm \quad (25)$$

4.2. Calculation of the tank wall thickness with regard to operating conditions

The nominal sheet thickness e_n is determined by the formula:

$$e_n = e + C + \delta_e + \delta_m \tag{26}$$

where: C – corrosion allowance,
δ_e – allowance for tolerances of rolled sheets (lower tolerance deviation),
δ_m – wall thickness allowance due to additional pressure stresses ($\delta_m = 0$) was assumed.

According to American data for the X6CrNiTi18-1 steel (marking according to AISI – 304), the reduction in wall thickness for a working water environment for a year is (with research being carried out for 15 years):

$$C_0 = 0.0074\ \mu m/year \tag{27}$$

Assuming $\tau = 15$ years as the expected tank life, the following was obtained:

$$C = C_0 * \tau = 0.0074 * 15 = 0{,}025\ \mu m \tag{28}$$

$$\delta_e = 0.14 \text{ mm according to EN10131:2006} \tag{29}$$

$$e_n = 1.92 + 0{,}14 + 0.00111 = 2.061\ mm \tag{30}$$

Due to the fact that the axial stresses are twice as small, the wall thickness considered in the axial direction will be smaller. The proposed solution assumes the thickness of the sheet as $e_n = 2$ mm.

For the proposed solution, the main stresses will be:

$$f_1 = \frac{P_d * D_m}{2 * e_n} = \frac{1{,}125 \frac{N}{mm^2} * 477{,}5\ mm}{2 * 2{,}5\ mm} = 106{,}43 \frac{N}{mm^2} = 106.43\ MPa \tag{31}$$

where: D_m – the average vessel diameter

$$f_2 = \frac{P_d * D_m}{4 * e_n} = 53.22\ MPa \tag{32}$$

$$f_{\text{rezdb}} = \sqrt{f_1^2 + f_2^2 - f_1 - f_2} = \sqrt{106{,}43^2 + 53{,}22^2 - 106{,}43 * 53{,}22} = 92.17\ MPa \tag{33}$$

According to (PN-EN 13445-1, 2014), the reduced stress should meet the condition:

$$f_{red} =\leq min\{\frac{5}{6}R_{P/T};\frac{1}{3}R_{M/T}\} \quad (34)$$

$$\frac{5}{6}R_{P/T} = \frac{5}{6} * 218 - 181{,}7\ MPa \quad (35)$$

$$\frac{1}{3}R_{M/T} = \frac{1}{3} * 495 = 165\ MPa \quad (36)$$

$$f_{rezdb} = 92.17\ MPa < 181.7\ MPa \text{ (the condition is met)} \quad (37)$$

Taking into account that the lowest acceptable stress value, which is significantly lower than experimentally determined, was used for calculations, the sheet thickness of e = 2 mm can be considered as sufficient.

5. EXPERIMENTAL RESEARCH

The main goal of the research was to obtain the data needed for the experimental assessment of the force of the vessel material at a certain value of its internal pressure load. The object undergoing testing was a pressure vessel with an average diameter of the cylindrical part of D_{sr} = 480 mm, made of X6CrNiTi18-10 steel. The thickness of the bottom walls and the cylinder itself is the same and equals 2 mm.

Fig. 5. The view of the vessel prepared for research

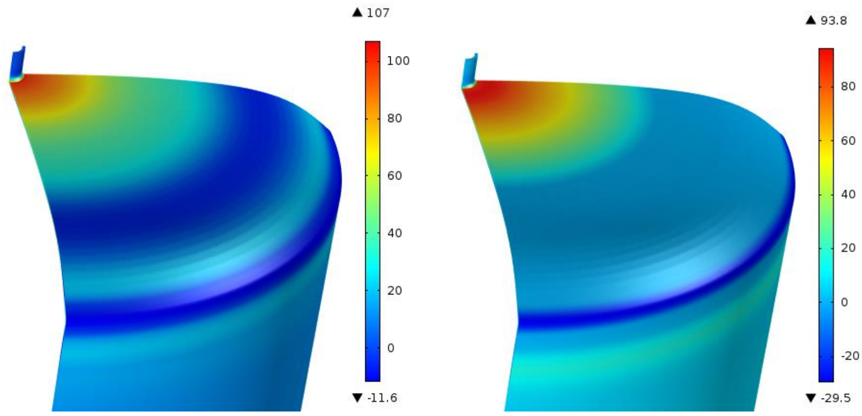


Fig. 6. Distribution of stresses in the bottom at a pressure of 0.3 MPa: circumferential stresses (left view) and radial stresses (right view)

Bottoms of cylindrical tanks usually have ellipsoidal or toroidal-spherical shapes and are characterized by their connection to the cylindrical surface of the tank in the axial plane of the cross-section at the place of the greatest curvature. This results in the disturbance of the membrane state of internal forces in the structure (Fig. 6). There is transverse force and a bending moment which cause changes in stress distribution, both in the bottom part and in the cylindrical part of the tank. Therefore, in order to select the most intensive areas of the tank, auxiliary calculations were made using the finite element method (Fig. 6).

The greatest stresses, both circumferential and longitudinal, occur in the region at the very bottom of the bottom. Then longitudinal stresses decrease, reaching negative values near the bottom connection with the cylindrical part. The circumferential stresses reach the second local maximum near the connection to the cylindrical part.

In order to determine the stresses at the characteristic points of the structure, strain gauges were glued on them. Figure 7 shows the location of the strain gauges on the bottom. For circumferential and longitudinal strain measurements, biaxial tensiometers of the type TF 3-2x / 120 from Tenmex were used.

Additionally, two strain gauges were installed on the cylindrical part. One was installed at a distance of 10 mm from the weld, and the second halfway up the cylindrical part of the tank. In addition, in order to verify the measurements, in the same configuration as for the top end, strain gauges were installed on the very bottom of the bottom. In order to compensate for the influence of temperature on the measurement results, strain gauges were mounted on the unloaded part of the tank.

For circumferential and longitudinal strain measurements, seven biaxial tensiometers of the type TF 3-2x / 120 from Tenmex were glued on to the head of the tank. In order to compensate for the influence of temperature on the unloaded steel X6CrNiTi18-10 plate single-axis strain gauges were mounted. Additionally, two strain gauges were installed on the cylindrical part. To measure the pressure in the tank, a piezoelectric transducer was used. The strain gauges and pressure transducer were integrated with the Catman measuring system using the four Hotminger Spider-8 amplifiers. During the test, data from 26 measurement channels were recorded. The pressure test was carried out as follows. After calibrating the measurement channels and stabilizing the indications of the strain gauges, the appropriate pressure in the tank was developed by the water pump.

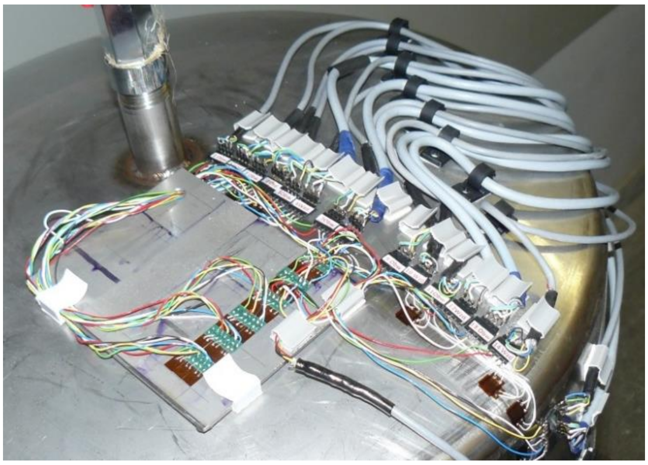


Fig. 7. View of tested tank with strain gauges

5.1. Pressure test

The pressure test was carried out as follows. After calibrating the measurement channels and stabilizing the indications from the strain gauges, the appropriate pressure value in the tank was forced by means of the pump and its constant value was maintained so as to obtain the determined conditions. This state was maintained for 15 seconds to be able to register a sufficiently large sample of measurement data. The measurement data was recorded and saved every 20 ms. At the same time, the tightness of the system was controlled, i.e., checks were made for the presence of water leaks. The pressure was measured in increments of 1 bar. The tests were carried out at a constant temperature of 24 °C.

In the pressure range from 1 to 9 bar, there was no slow pressure drop when pumping stopped. However, for higher pressures, its slow decline was noticeable, despite the pump's cessation. No leakage of water from the installation was found during this time. The pressure drop was caused by exceeding the yield point in the zone with the highest stresses, which caused the start of the “flow” of the material and the increase of the volume of the tank. This process deepened with increasing pressure, i.e. the gradient of pressure drop increased with increasing pressure level. This meant that an ever-larger material zone reached the limit of plasticity.

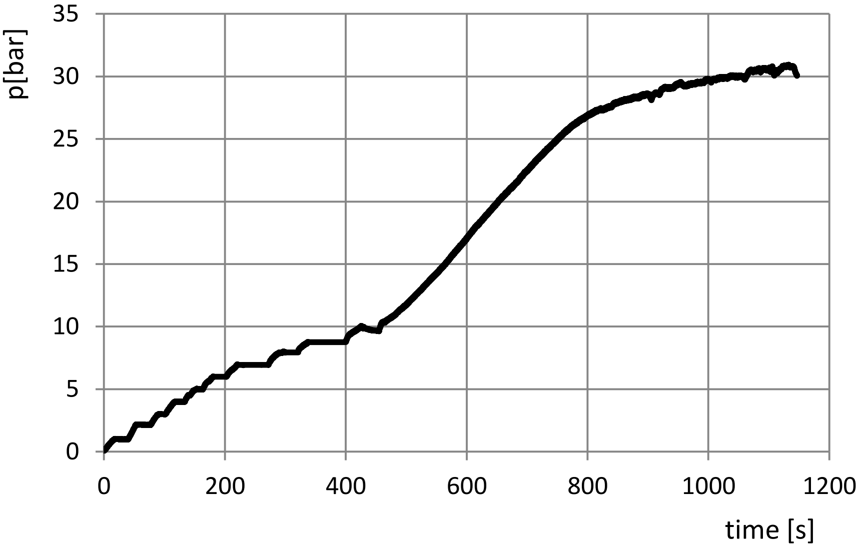


Fig. 8. Maximum pressure in the tank

The pressure test was carried out up to 30.9 bar. A further increase in pressure was not possible due to insufficient pump capacity, i.e. the plastic flow phenomenon was already so intense that it was not possible to achieve higher pressure values.

At the conclusion of the test, the tank was inspected and no leaks were found. All welded connections had been leak-proof and were undamaged.

Fig. 9. Changes in the shape of the tank bottom at an increase of pressure up to 30 bar

The degree of deformation, which is the effect of plastic deformation can be noticed by comparing the shape of the bottoms before and after the pressure test (Fig. 9).

5.2. Results of extensometer measurements

During the measurements, circumferential and longitudinal deformations were recorded, which in this case are the main deformations:

$$\epsilon_1 = \frac{1}{E} \cdot (\sigma_1 - \nu \cdot \sigma_2) \tag{38}$$

$$\epsilon_2 = \frac{1}{E} \cdot (\sigma_2 - \nu \cdot \sigma_1) \tag{39}$$

where: E – Young's modulus, ν – Poisson's ratio

In addition, dependencies are correct for the cylindrical part, based on membrane theory of coatings:

$$\sigma_1 = \frac{p \cdot D_{śr}}{2 \cdot e} \tag{40}$$

$$\sigma_2 = \frac{p \cdot D_{śr}}{4 \cdot e} \tag{41}$$

where: $D_{śr}$ – average diameter of the cylindrical part of the tank,
e – wall thickness.

Based on the above relationships and all of the following are measured: deformations, pressure, wall thickness and average tank diameter, it is possible to estimate the Young's modulus of elasticity and the Poisson's coefficient from the equations:

$$\epsilon_1 = p \cdot \frac{D_{śr}}{2e \cdot E} \cdot \left(1 - \frac{\nu}{2}\right) \tag{42}$$

$$\epsilon_2 = p \cdot \frac{D_{śr}}{2e \cdot E} \cdot \left(\frac{1}{2} - \nu\right) \tag{43}$$

The deformations ε_1 and ε_2 are linear functions of pressure. By specifying the directional coefficient of the functions (42) and (43), the Young's modulus and the Poisson's ratio values can be determined. The deformations ε_1 and ε_2 were recorded from the measurement point located on the cylindrical part of the tank, for averaged pressure values.

By approximating the measured measurement data ε_1 and ε_2 with linear functions in the pressure range 1 bar-6 bar, the directional coefficients, equations (42) and (43) were determined, which were respectively: $7.1077 \cdot 10^{-4}\ 1/MPa$ and $2.1407 \cdot 10^{-4}\ 1/MPa$.

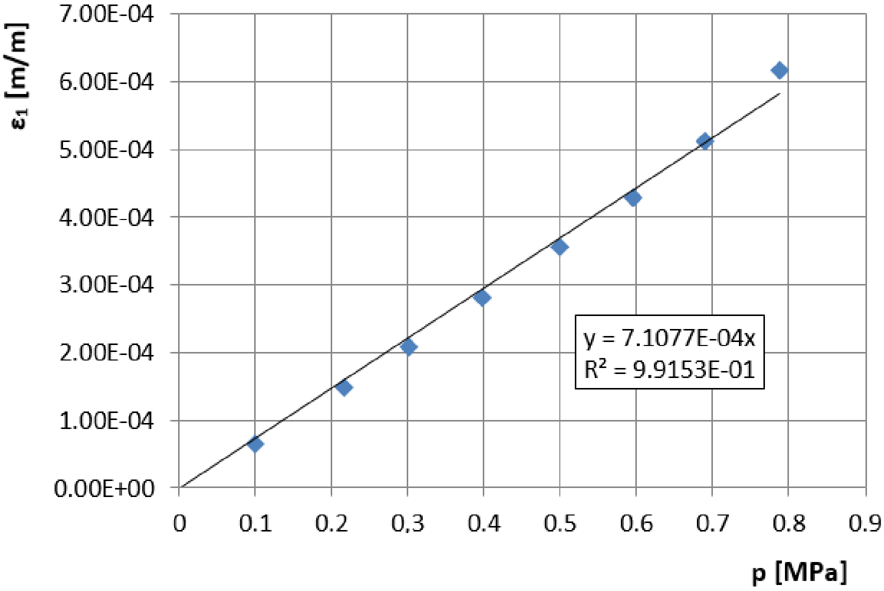


Fig. 10. Peripheral relative elongation – cylindrical part of the tank

On this basis, the Young's modulus and Poisson's ratio were estimated, the values of which were respectively: 150 GPa and 0.234.

Based on literaturę data, for the X6CrNiTi18-1 steel, Young's modulus at 20 °C ranges from 190 to 200 GPa. The Poisson's ratio varies from 0.24 to 0.3. The discrepancy between the literature data and the Young's modulus and the coefficient obtained from measurements was 23% and 13%, respectively. It is the result of measurement uncertainty of component quantities.

As can be seen, there is a linear dependence of strain at all measuring points in the pressure range from 1 to 8 bar, with the exception of point 6.

Using the determined Young's modulus and Poisson's ratio, the measured strains were transformed into stresses using the following relationships:

$$\sigma_1 = \frac{E}{1-\nu^2} \cdot (\varepsilon_1 + \nu \cdot \varepsilon_2) \tag{44}$$

$$\sigma_2 = \frac{E}{1-\nu^2} \cdot (\varepsilon_2 + \nu \cdot \varepsilon_1) \tag{45}$$

The circumferential stresses tend to decrease as the distance from the symmetry axis of the tank increases. This state maintains up to a point with a coordinate of approx. 230 mm, where the sign changes and the state of compressive stress changes.

At a distance of about 170 mm, circumferential (tensile) stresses reach the local minimum, then increase rapidly and change the sign.

The highest compressive (peripheral) stresses occur along the radius of the bottom passage into the cylindrical part. It is the effect of combining the ellipse with the straight part forming the cylindrical part of the tank. There is a transverse force and a bending moment that causes stress accumulation, both in the bottom part and in the cylindrical part of the tank.

Longitudinal stresses are positive (stretching) up to the point of 230 mm, where they change the sign. Along with the increase of the radius there is a slight decrease to the point with the coordinate of 70 mm. Then they grow again up to 220 mm, where they reach their maximum value.

As in the case of circumferential stresses, there is stress accumulation in the transition zone in the cylindrical part.

As you can see, this area has a decisive impact on the strength of the structure.

6. CONCLUSIONS

The conducted numerical and experimental studies show that the FEM method gives results inflated in relation to the analytical calculations, although it indicates the directions where the maximum stress increases should be expected.

The analytical method recommended by the relevant standards allows for the safe design of tanks and the maintenance of good safety ranges.

The material properties of steel given by standards are smaller than the actual values. When designing, for minimal safety factors, you must have your own material base, created as a result of tests in certified laboratories.

In the construction of tanks made of austenitic steel, exceeding the yield point is not synonymous with the destruction of the structure. For large stress values (exceeding the yield point) significant deformations and displacements occur, but the structure remains sealed.

Austenitic steels show good weldability, and the selection of a suitable welding technology ensures adequate weld strength and preservation of its integrity even when the parent material is destroyed.

REFERENCES

Bałon, P., & Świątoniowski, A. (2013). The influence of cold forming conception on the springback magnitude. *Key Engineering Materials*, *554–557*, 2299–2311.

Bałon, P., & Świątoniowski, A. (2014). Forming of automotive parts with nuts clinch process in comparison to welding of nuts. *Key Engineering Materials*, *611–612*, 1503–1510.

Bałon, P., & Świątoniowski, A. (2014). Stamping automotive parts with clinch nut process. *AIP Conference Proceedings*, *1567*, 971–974.

Bałon, P., & Świątoniowski, A. (2016a). Analiza procesu formowania dennicy hybrydowego urządzenia grzewczego. *Mechanik: miesięcznik naukowo-techniczny, 12,* 1840–1843.

Bałon, P., & Świątoniowski, A. (2016b). Improved method of springback compensation in metal forming analysis. *Strength of Materials*, *48*(4), 540–550.
Bałon, P., Świątoniowski, A., & Kiełbasa, B. (2017). The analysis of bottom forming process for hybrid heating device. *AIP Conference Proceedings*, *1896*, 170001.
Bałon, P., Świątoniowski, A., & Szostak, J. (2015). Łączenie elementów struktury samochodu z zastosowaniem wtłaczania i zgrzewania części złącznych. In J.Mucha (Ed.), *Badania i kierunki rozwoju technologii połączeń przetłaczanych na zimno: konstrukcje – procesy – zastosowanie* (pp. 221-236). Rzeszow, Poland: Oficyna Wydawnicza Politechniki Rzeszowskiej.
Bałon, P., Świątoniowski, A., Szostak, J., & Kiełbasa, B. (2016). The analysis of bottom forming process for hybrid heating device. In J. Mucha (Ed.), *Progressive technologies and materials* (pp. 7–22). Rzeszów, Poland: Oficyna Wydawnicza Politechniki.
Bałon, P., Świątoniowski, A., Szostak, J., & Kiełbasa, B. (2017). Springback compensation for a vehicle's steel body panel. *AIP Conference Proceedings*, *1896*, 080001.
Błachut, J., & Magnucki, K. (2008). Strength, stability, and optimization of pressure vessels: Review of selected problems. *Applied Mechanics Reviews*, *61*(6), 1–33.
Dyrektywa 97/23/WE Parlamentu Europejskiego i Rady z dnia 27.05.1997 w sprawie zbliżenia ustawodawstwa Państw Członkowskich dotycząca urządzeń ciśnieniowych. (1997).
Harvey, J. F. (2000). *Theory and Design of Pressure Vessels*. New Delhi: CBS Publishers & Distributors.
Krużelecki, J., & Proszowski, R. (2012). Shape optimization of thin-walled pressure vessel end closures. *Structural and Multidisciplinary Optimization*, *46*, 739–754.
Lakshmi Devi, E., & Hari Shankar, V. (2016). Modeling and Analysis of Cryogenic Pressure vessel using FEA. *International Journal of Engineering Trends and Technology (IJETT)*, *42*, (5), 261–265.
Lewiński, J., & Magnucki, K. (2010). *S*haping of a middle surface of a dished head of a circular cylindrical pressure vessel. *Journal of Theoretical and Applied Mechanics*, *48*(2), 297–307.
Lewiński, J., & Magnucki, K. (2012). Optimal shaping of middle surface of a dished head of circular cylindrical vessel with the help of B´ezier curve. *ASME 2010 Pressure Vessels and Piping Conference*, *5*, PVP2010-25208.
Li, J., Sheng, J., & Zhang, Y. (2012). Simulation Research of a type of Pressure Vessel under Complex Loading Part 2 Complex Load of the Numerical Analysis. *Advances in Intelligent Systems Research*, *26*, 4114.
PN-EN 13445-1. (2014). *Nieogrzewane płomieniem zbiorniki ciśnieniowe*. Część 1: Wymagania ogólne.
PN-EN 13445-3. (2014). *Nieogrzewane płomieniem zbiorniki ciśnieniowe*. Część 3: Projektowanie.
PN-EN 1993-1-6:2007. (2007). *Design of steel structures-Strength and stability of shell structures*.
PN-EN 1993-4-2:2009. (2009). *Design of steel structures-Tanks*.
Rozporządzenie Ministra Gospodarki z dnia 21.12.2005 w sprawie zasadniczych wymagań dla urządzeń ciśnieniowych i zespołów urządzeń ciśnieniowych. (2005).
Rozporządzenie ministra Gospodarki, Pracy i Polityki Społecznej z dnia 9.07.2003 w sprawie warunków technicznych dozoru technicznego w zakresie eksploatacji niektórych urządzeń ciśnieniowych. (2003).
Rozporządzenie Rady Ministrów z dnia 16.07.2002 w sprawie rodzajów urządzeń technicznych podlegających dozorowi technicznemu. (2002).
Ventsel, E., & Krauthammer, T. (2001). *Thin Plates and Shells. Theory, Analysis and Applications*. Basel, New York, USA: Marcel Dekker Inc.
Warunki Urzędu Dozoru Technicznego WUDT/UC/2003 (2005). *Urządzenia ciśnieniowe*. Wydanie II, Warsaw.
Ziółko, J. (1986). *Metal Tanks for Liquids and Gases* (in Polish). Warsaw, Poland: Arkady.

Applied Computer Science, vol. 14, no. 4, pp. 54–69
doi:10.23743/acs-2018-29

Submitted: 2018-10-24
Revised: 2018-12-07
Accepted: 2018-12-19

inertia moment, induction motor, wavelet transformation, backlash zone, neural network weights

Marcin TOMCZYK[*], *Barbara BOROWIK* [0000-0002-0197-4902][**], *Mariusz MIKULSKI*[***]

IDENTIFICATION OF A BACKLASH ZONE IN AN ELECTROMECHANICAL SYSTEM CONTAINING CHANGES OF A MASS INERTIA MOMENT BASED ON A WAVELET–NEURAL METHOD

Abstract

In this article a new method of identification of a backlash zone width in a structure of an electromechanical system has been presented. The results of many simulations in a tested model of a complex electromechanical system have been taken while changing a value of a reduced masses inertia moment on a shaft of an induction motor drive. A wavelet analysis of tested signals and analysis of weights that have been obtained during a neural network supervised learning - have been applied in a diagnostic algorithm. The proposed algorithm of detection of backlash zone width, represents effective diagnostic method of a system at changing dynamic conditions, occurring also as a result of mass inertia moment changes.

[*] Electrical School No. 1 in Krakow them. Silesian Insurgents, Kamieńskiego 49 Street, 30-644 Kraków, Poland, e-mail: tomczykmarcin@poczta.fm
[**] Cracow Univeristy of Technology, Faculty of Physics, Mathematics and Computer Science, Institute of Computer Science, Warszawska 24 Street, 31-155 Kraków, Poland, e-mail: bborowik@pk.edu.pl,
[***] State University of Applied Sciences in Nowy Sącz, Institute of Technology, Zamenhofa 1a Street, 33-300 Nowy Sącz, Poland, e-mail: mmikulsk1@poczta.onet.pl

1. INTRODUCTION

In this paper it has been assumed, that information has been available by measuring signals, and this most often is the base of the diagnostic reasoning methods in electromechanical systems. The process of simultaneous classification of signals in the field of time and frequency is possible due to the use of transformation methods, which allow to test these signals spectral properties. Tested signal can be presented as linear combination of certain base orthogonal signals. This method minimizes the signal model. For the purpose of the minimization of the set of important coefficients of a distribution it is necessary to fit shapes of base functions to the analyzed signal (Zając, 2009).

A progress in technology of signal processing has caused that during the last few years a big number of induction motor drives diagnostics methods by means of wavelet analysis have been studied and presented in numerous papers. Within the last two years more and more scientific works have occurred that have been using time-frequency methods and neural networks in fault diagnostics of industrial structures.

During the last few years there have been published some works that have been using a combination of a wavelet transformation with an artificial intelligence method, and which have been presented as faults detection techniques. It is worthy to mention some of them, like the following ones:

- using a combination of a neural network and a wavelet transformation for diagnostics purposes for the induction motor drive (Kowalski, 2005),
- analysis of descriptions of a current of a motor (MCSA) for a stator current with the use of a wavelet analysis in faults detection of the broken rotor’s barrier in the transitional area (Douglas, Pillay & Ziarani, 2003),
- presentation of methods of detection of rotor’s breakage in the cage induction motor on the basis of analysis of the stator’s current by means of using a Fourier transformation and a discrete wavelet transformation (Da Costa, Kashiwagi & Mathias, 2015),
- analysis of stator’s current using the combination of a discrete wavelet transformation and a feedforward neural network and a neural network with radial activation functions (Sridhar, Uma Rao & Jade, 2016), as well as analysis of vibration - for the purpose of faults diagnostics of the induction motor drive using a wavelet decomposition and methods of clustering data FCM based on fuzzy logic (Chandralekha & Yayanthi, 2016),
- presentation of a possibility of a mechanical estimation of state’s variables of a two-mass drive with the use of a perceptron neural network MLP (*Multi Layer Perceptron*) (Orłowska-Kowalska & Szabat, 2007),
- diagnostics of the induction motor using a packet wavelet transformation for detection of faults of a stator’s current (Annamalai & Swaminathan, 2016).

2. PRESENTATION OF THE METHODOLOGY AND RESEARCH ON FAULT IDENTIFICATION ALGORITHM

Diagnostic tests have been carried out using a working machine in the form of a dynamic mass-absorbing-resilient element connected to the load of the induction motor. There have been accepted nominal conditions of the induction motor, and this model has been built in a stationary coordinate system related to the stator (model α, β, 0). The tested backlash zone width occurs between the rod of the induction motor drive and a working machine drive wheel.

A simplified form of the connection of the induction motor with a working machine has been shown on a diagram in Fig. 1.

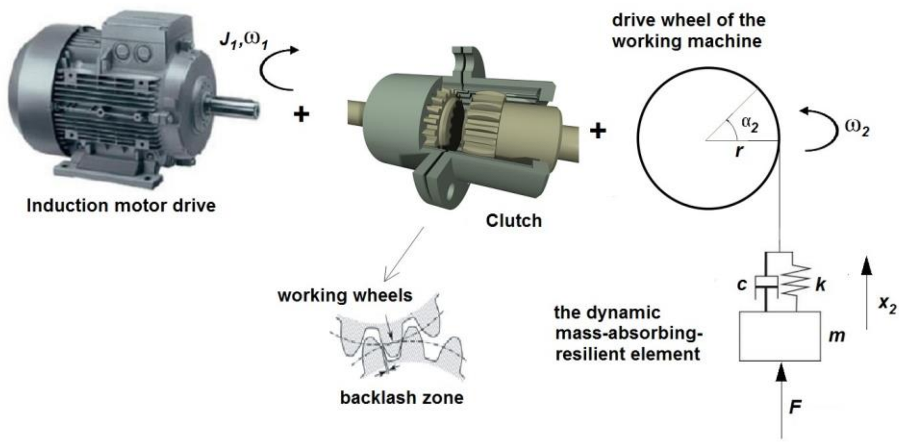


Fig. 1. A diagram showing the connection of the induction motor drive with the dynamic mass-absorbing-resilient element by means of a clutch, in which a backlash zone occurs and gradually increases

The following parameters of the induction motor have been applied in the conducted simulation tests (parameters of a substitute circuit are expressed in relative units): circuit stator relative resistance r_s = 0.059 [Ω], circuit rotor relative resistance r_w = 0.048 [Ω], relative reactance of the dispersion circuit stator x_s = 1.92 [Ω], relative reactance of the dispersion circuit rotor x_w= 1.92 [Ω], relative reactance of the dispersed circuit x_m= 1.82 [Ω], mechanical time constant T_m = 0.86 [s].

All tests have been carried out within a MATLAB / Simulink R2017a environment.

3. IDENTIFICATION TESTS OF THE BACKLASH ZONE WIDTH IN AN ELECTROMECHANICAL SYSTEM CONTAINING A VISCOUS FRICTION

The tests have been carried out in four test groups, with the following four different values of apparent viscosity coefficient c: 0.8 [N·s/m], 1.0 [N·s/m], 1.25 [N·s/m] and 1.4 [N·s/m]. Each test group has contained seven cases with different inertia moment J_1 values. Results of the simulations for all these physical quantities have been written in the matrix $N_{[7,2048]}$. Elements of the matrix N have been written for each coefficient value of the viscosity friction c. The value of an inertia moment J_1 has been determined as the percentage in relation to its nominal value J_1, down with value A% and up with value C%. Formal changes of the inertia moment J_1 have been written in matrix K_1 in the following order: K_1 = [nominal value of the inertia moment (J_1 = 0.87 [kg*m^2]), A=2.5% (J_1 = 0.8482 [kg*m^2]), A = 5% (J_1 = 0.8265 [kg*m^2]), A = 7.5% (J_1 = 0.8047 [kg*m^2]), A = 12.5% (J_1 = 0.7612 [kg*m^2]), C = 2.5 % (J_1 = 0.8917 [kg*m^2]), C = 5% (J_1 = 0.9135 [kg*m^2])]. In all executed simulation tests the inertia moment J_2 of the drive wheel of a working machine has been assumed to have value equal to 0.348 [kg*m^2].

For each of the seven cases of changes of an inertia moment J_1 there have been carried out simulation tests for six backlash zone widths. The backlash zone width values have been taken in sequence from the epsilon matrix, in the following order: K = [0.0025, 0.00375, 0.005, 0.0075, 0.009, 0.01]. All tests have been carried out for the discontinuity in zero value c_1 = 0.2.

The wavelet type and its order they have been selected in such a way that the shape of the basic wavelet approximately would be adequate to the character of the transient course of the tested physical quantity that would be obtained as a result of a simulation for the case of the smallest backlash value. Based on the carried out tests the following selections of wavelets have been made for individual physical variables, with decomposition level 10:

a) a linear acceleration of the induction motor drive a_1 – wavelet function symlet of the order 5,
b) an electromagnetic moment of the induction motor drive m_{el} – wavelet function daubechies of the order 6,
c) an angular speed of the induction motor drive rotor ω_1 – wavelet function symlet of the order 5,
d) a linear acceleration of a mass a_2 – wavelet function daubechies of the order 6,
e) a linear speed of a mass v_2 – wavelet function symlet of the order 5.

In these simulation tests, it has been assumed that the process of the electromechanical system dynamics testing within the backlash zone starts when the expression specified in the left part of the following inequality (1) is smaller than its right part:

$$|\alpha_1 - \alpha_2| < \frac{K_{(i)}}{r}; \quad i \in \{1, 2, \ldots 6\}, \tag{1}$$

where: r – a radius of the drive wheel of a working machine [m],
$K_{(i)}$ – value that has been taken sequentially from the matrix K, corresponding to the given backlash value of its mechanical connection,
i – index number within the matrix K.

The location change angle α_1 for rod masses of the induction motor drive [rad] has been calculated using the formula: $\alpha_1 = \int \omega_1 \cdot dt$, and the location change angle α_2 of the drive wheel of the working machine [rad], has been calculated using the formula: $\alpha_2 = \int \omega_2 \cdot dt$, where ω_1 is the angular speed, done by rod masses of the induction motor drive [rad/s], and ω_2 is the angular speed done by a drive wheel of the working machine [rad/s].

After satisfying the condition determined by the inequality described by formula (1) the electromagnetic moment of the induction motor drive in the tested electromechanical system is set to zero value.

While conducting these simulations it has been assumed that specified earlier matrix K would contain 2000 samples of a signal, chosen in a sequence starting from the time obtaining in the tested model of an electromechanical system a backlash zone and also 48 samples preceding this point in time. This assumption has been presented in Fig. 2.

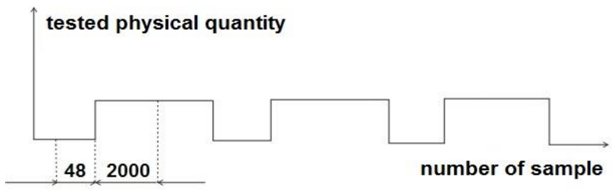


Fig. 2. Testing dynamics in a backlash zone for all physical quantities carried out by means of selected time range choices (samples of the tested signal)

3.1. Description of the learning process of the three-layer backpropagation neural network

During simulations discussed in the Section 3.2 the value of the apparent viscosity coefficient *c* has been set up to 1.0, the value of the backlash zone width has been allowed to change and the value of the inertia moment J_1 has been set up, and all these assumptions have had impact on all values of variables and matrices, that have been calculated or obtained during the conducted testing.

Nonlinear neural network's learning algorithm based on the backpropagation of errors method enables reproduction of the assumed values of errors into deeper layers of this network (to which a direct access is not available), by back propagating errors detected on this network's output. During analysis of a single hidden layer neuron – errors of all of neurons, to which this single neuron has transferred his output signal have been taken into account. Then these errors have been summed up taking also into consideration these neurons weights. Fig. 3 presents used in the simulations 3-layers backpropagation neural network.

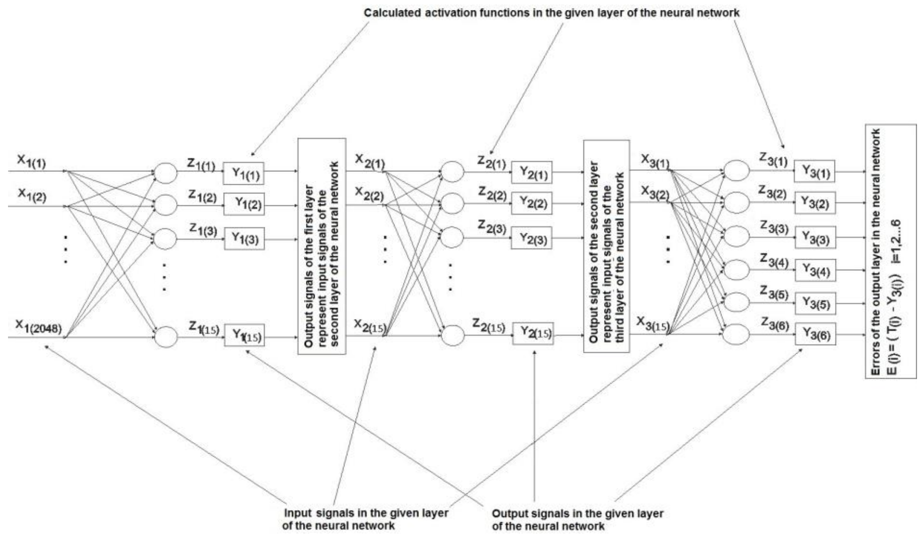


Fig. 3. Used in the simulations 3-layers backpropagation neural network

The first layer of this network represents the given set of sample pairs of this neural network input signals X_1 and values of weights W_1. The second and third layer of this network represents the set of the output signals of the second and third neural network layer X_2 and X_3, as well as sample values of their weights (i.e. W_2 and W_3).

The identification procedure could be successfully carried out in conditions of changing values of the reduced masses inertia moment and the induction motor drive J_1, which has been stiffly connected to the rotor.

The results of the backpropagation neural network processing depend on the appropriate defining of the input signals in the first layer X_1 and also on the initial selection of the weight values in the first, second and the third layer (i.e. in W_1, in W_2 and in W_3).

Input values X_1 in the first layer of the neural network represent values of the matrix N normalized to the range $[h_2,k_2]$ and have been calculated according to the formula:

$$X_{1(p)(j)} = \left(\frac{\left(N_{(p)(j)} - k_1\right)}{h_1 - k_1} \right) \cdot \left(h_2 - k_2\right) + k_2; \; j \in \{1,2,\ldots,2048\}, \; p \in \{1,\ldots,7\}, \tag{2}$$

where: N – values of the matrix registered during the test,
h_1 and k_1 – the minimal and the maximal value of the matrix N determined during the test,
h_2 and k_2 – the initial and the final value of the range, which contains normalized values of the matrix N calculated during the test,
p – a column number of the matrix "K_4".

For all tested physical quantities variables h_2 and k_2 represent the minimal (h_2 = min(h_3, h_4)) and the maximal (k_2 = max(h_3, h_4)) value. The values of the variables h_3 and h_4 have been calculated in the following way: $h_3 = (h_5 - h_6)$ and $h_4 = (h_6 - h_5)$, where h_5 and h_6 have been calculated as follows:

$$h_5 = \left(N_{1(1)} - N_{1(2)}\right) - \left(N_{1(2)} - N_{1(3)}\right) - \left(N_{1(3)} - N_{1(4)}\right) - \left(N_{1(4)} - N_{1(5)}\right), \tag{3}$$

$$h_6 = \left(N_{2(1)} - N_{2(2)}\right) - \left(N_{2(2)} - N_{2(3)}\right) - \left(N_{2(3)} - N_{2(4)}\right) - \left(N_{2(4)} - N_{2(5)}\right). \tag{4}$$

Values of two matrices N_m for $m \in \{1, 2\}$ represent arithmetic means of the matrix N and during the test values of their rows' elements have been calculated for a given value of the inertia moment J_1 using the following formula:

$$N_{m(p)(i)} = \frac{\sum_{j=1}^{n} N_{(p)(j)}}{n},$$

for $i \in \{1, ..., 5\}$, $p \in \{1, ..., 7\}$ and for $m \in \{1, 2\}$; and with the notice that for $m = 1$ (i.e. for matrix N_1): for $i = 1$ n is equal to *100;* for $i = 2$ n is equal to *200;* for $i = 3$ n is equal to 300*;* for $i = 4$ n is equal to 400 and for $i = 5$ n is equal to 500*;* and for $m = 2$ (i.e. for matrix N_2) with the notice that: for $i \in \{1, ..., 5\}$ and $p \in \{1, ..., 7\}$*;* and with the notice that for $i = 1$ n is equal to 256*;* for $i = 2$ n is equal to 512*;* for $i = 3$ n is equal to 768*;* for $i = 4$ n is equal to 1024 and for $i = 5$ n is equal to 1280.

The best result of calculations has been obtained while using the two cases of determining the values of variables h_3 and h_4 for simulations with apparent viscosity coefficient coefficient $c = 1.0$.

During the test calculations of values of variables h_3 and h_4 have been carried out for the specified earlier value of inertia moment J_1 in the following way:

$$h_3 = \begin{cases} (h_5 - h_6), & for \;\; h_7 <= h_8 \\ \begin{cases} \dfrac{(m_1)}{(h_{11})}, & for \;\; h_9 <= h_{10} \;\; and \;\; h_7 > h_8 \\ h_{12}, & for \;\; h_9 > h_{10} \end{cases} \end{cases},$$

$$h_4 = \begin{cases} (h_6 - h_5); & for \;\; h_7 <= h_8 \\ \begin{cases} (m_1)\cdot(h_{11}); & for \;\; h_9 <= h_{10} \;\; and \;\; h_7 > h_8 \\ m_2; & for \;\; h_9 > h_{10} \end{cases} \end{cases}, \tag{5}$$

where: h_7, h_9 – values of the variables representing the result of calculations executed with using specific, appropriate variables during the test,
h_8 the minimal value of the matrix N_3 determined during the test,
h_{10}, h_{11}, h_{b12} – values of these variables have been obtained in the test,
m_1,m_2 – arithmetic means calculated during the test.

Value of the variable h_7 has been calculated in the following way: $h_7 = h_{13} + (m_3 - h_8)$. The variable h_{13} carries the maximum value of the matrix N_3 (i.e. $h_{13} = \max(N_{3(p)(i)})$, for $i \in \{1, ..., 6\}$ and $p \in \{1, ..., 7\}$). The variable h_8 carries the minimum value of the matrix N_3 (i.e. $h_8 = \max(N_{3(p)(i)})$ for $i \in \{1, ..., 6\}$ and $p \in \{1, ..., 7\}$). The variable m_3 is the arithmetic mean calculated for the set up number of elements of the matrix N_3 during the test – i.e.

$$m_3 = \frac{\sum_{i=1}^{3} N_{3(p)(i)}}{\int 3},$$

for $i \in \{1, ..., 6\}$ and $p \in \{1, ..., 7\}$).

Elements of the matrix N_3 – carry the results of the cyclic summations of values of variable h_{16} calculated for the tests: $N_{3(\mathrm{p})(i)} = \sum h_{16}$ for $N_{(p)(i,j)} > \partial_{1(p)(i)}$, for $i \in \{1, ..., 6\}$, j $\in \{1, ..., 2048\}$, and for p $\in \{1, ..., 7\}$.

$\partial_{1(p)(i)}$ is the standard deviation of elements of matrix N i-th row: $\partial_{1(p)(i)} = \sqrt{\frac{\left(N_{(p)(i,j)} - m_{4(i)}\right)^2}{2048}}$, for $i \in \{1, ..., 6\}$, $j \in \{1, ..., 2048\}$, and $p \in \{1, ..., 7\}$.

The value of the arithmetic mean m_4 has been calculated using the formula:

$$m_{4(\mathrm{p})(i)} = \frac{\sum_{j=1}^{2048} N_{(p)(i,j)}}{2048},$$ for $i \in \{1, ..., 6\}$, $j \in \{1, ..., 2048\}$, and $p \in \{1, ..., 7\}$.

The variable h_{16} stores the biggest value of the matrix N, which has been registered during simulations (i.e. h_{16} = max($N_{3(p)(i,j)}$), for $i \in \{1, ..., 6\}$, $j \in \{1, ..., 2048\}$ and $p \in \{1, ..., 7\}$).

The values of variables h_{b9}, h_{b10}, h_{14} and h_{15} have been calculated as follows:

$h_9 = | h_{14} - h_{18} |$ and h_{10} = max($h_{15,}$ h_{16}), $h_{14} = \sum h_{16}$, for $N_{(p)(j)} > m_{2(p)}$, for $j \in \{1, ..., 2048\}$, and $p \in \{1, ..., 7\}$; $h_{15} = | h_{17} - h_{18} |$. The value of the variable m_2, which is arithmetic mean of the tested matrix N, has been calculated as: $m_{2(\mathrm{p})} = \frac{\sum_{j=1}^{2048} N_{(p)(j)}}{2048}$, for $j \in \{1, ..., 2048\}$, and $p \in \{1, ..., 7\}$. The variables h_{15} and h_{16} are obtained and used during the test. The variable h_{17} and h_{18} store the biggest and the smallest value of the matrix N_4 (i.e. $h_{17} = \max\left(N_{4(p)(i)}\right)$ and $h_{18} = \min\left(N_{4(p)(i)}\right)$, for $i \in \{3,4, ..., 6\}$ and $p \in \{1, ..., 7\}$).

The variable h_1 stores the smallest value determined for selected elements of the matrix N_4.

The matrix N_4 has been calculated using the formula: $N_{4(\mathrm{p})(i)} = \sum h_{16}$, for $N_{(p)(i,j)} > m_{4(p)(i)}$, for indexes $i \in \{1, 2, ..., 6\}$, $j \in \{1, 2, ..., 2048\}$ and $p \in \{1, ..., 7\}$. The variable $m_{1(p)}$ is the arithmetic mean calculated for the specific row of the matrix N (i.e. $m_{1(\mathrm{p})} = \frac{\sum_{j=1}^{2048} N_{(p)(1,j)}}{2048}$, for $j \in \{1, ..., 2048\}$ and $p \in \{1, ..., 7\}$), where N is the matrix registered during the test, while the backlash zone width has its smallest value.

The value of the variable h_{11} has been calculated in the test using the following formula: $h_{11} = (h_{19} - h_{21}) - (h_{20} - h_{21}) - (m_5 - h_{21})$,

where: h_{19}, h_{20} – variables storing the smallest and the biggest value of the matrix N_5,
h_{21} – a variable used in the test,
m_5 – the value of the arithmetic mean of the matrix N_5.

The values of the matrix N_5 have been calculated during the test in the following way:

$$N_{5(p)(i)} = (N_{2(p)(i,1)} - N_{2(p)(i,2)}) - (N_{2(p)(i,2)} - N_{2(p)(i,3)}) - (N_{2(p)(i,3)} - N_{2(p)(i,4)}) - (N_{2(p)(i,4)} - N_{2(p)(i,5)}),$$

for $i \in \{1,2, ..., 6\}$ and $p \in \{1, ..., 7\}$.

The variables h_{19} and h_{20} store the smallest and the biggest values of the matrix N_5 (i.e. $h_{19} = \min(N_{5(p)(i)})$ and $h_{20} = \max(N_{5(p)(i)})$, for $i \in \{1, ..., 6\}$ and $p \in \{1, ..., 7\}$). The value of the arithmetic mean m_5 has been calculated using the formula: $m_5 = \frac{\sum_{i=1}^{6} N_{5(p)(i)}}{6}$ for $p \in \{1, ..., 7\}$. The value of the variable h_{21} has been calculated in the following way:

$$h_{21} = (N_{1(1)} - N_{1(2)}) + (N_{1(2)} - N_{1(3)}) + (N_{1(3)} - N_{1(4)}) + (N_{1(4)} - N_{1(5)}).$$

The value of the variable h_{12} has been calculated according to the following summation: $h_{12} = \sum h_{16}$, for $N_{(p)(j)} > \partial_{2(p)}$, for $j \in \{1,2, ..., 2048\}$ and $p \in \{1, ..., 7\}$, and $\partial_{2(p)}$ is the standard deviation calculated as: $\partial_{2(p)} = \sqrt{\frac{(N_{(p)(j)} - m_2)^2}{2048}}$, for $j \in \{1, ..., 2048\}$ and $p \in \{1, ..., 7\}$. The variable m_2 is the arithmetic mean of the tested matrix N.

Initial values of the weights in all layers of the backpropagation neural network have been calculated using the arithmetic mean m_1 described earlier (i.e. $W_{(d)(i,j)} = m_{1(\mathrm{i})}$, with the notice that for $d = 1$ $i \in \{1, ..., 15\}$ and $j \in \{1, ..., 2048\}$; for $d=2$ $i \in \{1, ..., 15\}$ and $j \in \{1, ..., 15\}$; for $d=3$ $i \in \{1, ..., 6\}$ and $j \in \{1, ..., 15\}$, and where: m_1 – is the value of the arithmetic mean calculated in the determined group of analysis, d – the layer number of the neural network.

The output signals Z_1 in the first layer of the neural network have been calculated using the formula: $Z_{1(i)} = \sum_{j=1}^{2048} X_{1(j)} \cdot W_{1(i,j)}$, with the notice that for $d \in \{1, 2\}$ the index i belongs to the set {1, ..., 15}; and for d=3 the index i belongs to the set {1, ..., 6}. The output signals of each layer have been used to calculate the value of the activation function of neurons according to the formula: $Y_{(d)(i)} = \frac{1}{1+\exp^{\left(-Z_{(d)(i)}\right)}}$, with the same notice as mentioned before, i.e. for $d \in \{1,2\}$ the index i belongs to the set {1, ..., 15}; and for d=3 the index i belongs to the set {1, ..., 6}. The input signals in the hidden layer, as well as in the output layer of the neural network have been calculated using the formula: $X_{(d)(j)} = Y_{(d-1)(i)}$, for $d \in \{2,3\}$, $i \in \{1,...15\}$, and $j \in \{1,...15\}$. The output signals Z_2 and Z_3 from neurons within the hidden layer and the output layer of the neural network have been calculated using the formula: $Z_{(d)(i)} = \sum_{j=1}^{15} X_{(d)(j)} \cdot W_{(d)(i,j)}$, with the notice that for d = 2 the index i belongs to the set {1, ..., 15}; and for d = 3 the index i belongs to the set {1, ..., 6}. The general formula of the output of the exponential activation function for the given input to the neural network T is the following: $T_{(i)} = \exp(-h_2)$, for $i \in \{1,...6\}$ (h_2 – was explained at the bottom of the 6th page of this article). The index i shows the row number within the matrix T. Values of errors E on neurons in the output layer have been calculated using the formula: $E_{(i)} = T_{(i)} - Y_{3(i)}$, for $i \in \{1,...6\}$. The value of the root-mean-square error RMS_1 has been determined according to the formula:

$$RMS_1 = \sqrt{\frac{\sum_{i=1}^{6} \left(E_{(i)}\right)^2}{6}}$$

It has been assumed that the process of learning of this neural network would end, when the following condition would be satisfied: $RMS_1 < \delta_1$, where δ_1 is the experimentally assumed value during conducted simulations for the necessity of stopping the neural network learning process.

The correction of weights, carried out in the respective layers of this neural network has been calculated according to its general formula (Osowski, 1996): $W_{(d)(i,j)} = W_{(d)(i,j)} + l \cdot \sigma_{(d)(i)} \cdot X_{b(d)(j)}$, with the notice that for d = 1 $i \in \{1, ..., 15\}$ the index $j \in \{1, ..., 2048\}$; for d = 2 the index $i \in \{1, ..., 15\}$ and the index $j \in \{1, ..., 15\}$; for d = 3 the index $i \in \{1, ..., 6\}$ and the index $j \in \{1, ..., 15\}$, where: l – the learning coefficient of the neural network, $l \in [0, ..., 1]$,

$\sigma_{(d)(i)}$ – the error value calculated in a given layer using the formula:

$\sigma_{(d)(i)} = E_{1(d)(i)} \cdot \sum_{i=1}^{15} W_{(d+1)(i,j)} \sigma_{(d+1)(i)}$, with the notice that for d = 1 the indexes i and j belong to the set {1, ..., 15}; and for d = 2 the index i belongs to the set {1, ..., 15}, but the index j belongs to the set {1, ..., 6}. For d = 3 and for the index i belonging to the set {1, ..., 6} the error value in the third layer has been calculated using the formula: $\sigma_{(d)(i)} = E_{1(d)(i)} * \left(T_{(\mathrm{d})(i)} - Y_{3(\mathrm{d})(i)}\right)$. $E_{(d)(i)}$ – is the value of the derivative of the sigmoidal activation function, calculated on the i-th neuron within the layer d of the neural network T, using the formula: $E_{1(d)(i)} = Y_{(d)(i)} \cdot \left(1 - Y_{(d)(i)}\right)$, with the notice that for $d \in$ {1, 2} the index i belongs to the set {1, ..., 15}; and for d = 3 the index i belongs to the set {1, ..., 6}.

Pattern matrices W_w have been created for all tested physical quantities for the given group of tests. It has been noticed that creation of pattern matrices and tested matrices W_t with using the values of the W_3 neural network weights matrix has allowed to obtain much better identification of the backlash zone width, compared to using weights matrix from other layers of this neural network. This has been noticed for all tested physical quantities for the given group of tests.

Matrix W_4 represents values of weights of the first neuron of the output layer W_3, i.e.: $W_{4(j)} = W_{3(1,j)}$ for $j \in$ {1, ..., 15}. The values of the pattern matrix W_w as well as the values of the tested matrix W_t, which have been applied in the process of identification of the backlash zone width have been obtained using the following formulas: $W_{w(i,j)} = W_{4(i,j)}$, for $i \in$ {1, ..., 6} and $j \in$ {1, ..., 15}; and $W_{t(j)} = W_{4(j)}$, for $j \in$ {1, ..., 15}. Then it has been created the matrix D in the following way: $D_{(i)} = \sum_{j=1}^{15} \left| W_{t(j)} - W_{w(i,j)} \right|$ for $i \in$ {1, ..., 6}. Then, based on calculating the minimal value of the matrix D it has been determined the searched index nr_5, $nr_5 = i$, such that $D_{(nr_5)} = \min\left(D_{(i)}\right)$. The searched index i (i.e. called nr_5) in the matrix D determines the column number in the matrix K, which contains the correct backlash zone width.

3.2. The results of the simulations of the diagnostic algorithm applied for identification of a backlash zone width in electromechanical system

In the column named *The results* there have been presented assumed in the fault identification process of the backlash zones widths. In this column using bold font there have been distinguished obtained the final values of the matrix *D*, resulting from applying the identification procedure, which has allowed to obtain the fault identification of number. The tables 1 through 4 show all the values of all parameters that have been used in the tests. It could be noticed that for the neural network learning coefficient *l* changing within the [0.6...0.9] range and for the value of the variable δ_1, used for stopping the neural network learning process changing within the [0.07...0.08] range – for all analyzed physical quantities have been obtained satisfactory result concluded with finding the minimum value of the matrix *D*, which was determining the searched index nr_5 (i.e. the column number in the matrix *K*, which contains the correct backlash zone width).

Tab. 1. Exemplified in matrix *D* the results of tests for linear acceleration of the induction motor drive a_1

Applied in the test parameters	The results	Applied in the test parameters	The results
backlash zone = 0.0075, inertia moment J_1 = 0.8917, c = 1.5, epochs = 11, l = 0.9, δ_1 = 0.08	0.2736 0.2727 0.1933 **0.0006** 0.0116 0.0791	backlash zone = 0.0075, inertia moment J_1 = 0.8917, c = 1.5, epochs = 12, l = 0.9, δ_1 = 0.07	0.2693 0.2684 0.1864 **0.0006** 0.0122 0.0824
backlash zone = 0.0025, inertia moment J_1 = 0.8482, c = 0.8, epochs = 12, l = 0.9, δ_1 = 0.08	**0.0020** 0.1912 0.2786 0.3581 0.2390 0.0238	backlash zone = 0.0025, inertia moment J_1 = 0.8482, c = 0.8, epochs = 15, l = 0.8, δ_1 = 0.07	**0.0021** 0.1733 0.3289 0.4099 0.2226 0.0262

Tab. 2. Exemplified in matrix *D* the results of tests for electromagnetic moment of the induction motor drive m_{el}

Applied in the test parameters	The results	Applied in the test parameters	The results
backlash zone = 0.009, inertia moment J_1 = 0.8265, c = 1.5, epochs = 22, l = 0.9, δ_1 = 0.08	0.0616 0.0106 0.0081 0.0082 **0.0004** 0.0068	backlash zone = 0.009, inertia moment J_1= 0.8265, c = 1.5, epochs = 26, l = 0.9, δ_1 = 0.07	0.0172 0.0466 0.0087 0.0088 **0.0004** 0.0073
backlash zone = 0.0025, inertia moment J_1 = 0.9135, c = 0.8, epochs = 19, l = 0.9, δ_1 = 0.08	**0.0038** 0.2669 0.2659 0.2584 0.2514 0.2501	backlash zone = 0.0025, inertia moment J_1 = 0.9135, c = 0.8, epochs = 25, l = 0.8, δ_1 = 0.07	**0.0041** 0.3013 0.3002 0.2921 0.2528 0.2514

Additionally, as a result of decreasing a previously set up range of the variable δ_1 , used for stopping the neural network learning process to the range from 0.08 to 0.07, by lowering the range of the neural network learning coefficient l to the range from 0.9 to 0.6, it has been noticed a big increase in number of epochs of the neural network learning process. All this is visible in all four tables, that present the results of the analysis.

Tab. 3. Exemplified in matrix *D* the results of tests for angular speed of the rotor of the induction motor drive ω_1

Applied in the test parameters	The results $\cdot 10^{-3}$	Applied in the test parameters	The results $\cdot 10^{-3}$
backlash zone = 0.0025, inertia moment J_1 = 0.8047, c = 0.8, epochs = 22, l = 0.9, δ_1 = 0.08	0.3001 0.1935 **0.0185** 0.2341 0.4384 1.2073	backlash zone = 0.0025, inertia moment J_1 = 0.8047, c = 0.8, epochs = 25, l = 0.8, δ_1 = 0.08	0.3008 0.1939 **0.0185** 0.2346 0.4393 1.2098
backlash zone = 0.01, inertia moment J_1 = 0.87, c = 1.25, epochs = 22, l = 0.9, δ_1 = 0.08	0.2754 0.3908 0.1967 0.0907 0.1643 **0.0244**	backlash zone = 0.01, inertia moment J_1 = 0.87, c = 1.25, epochs = 39, l = 0.6, δ_1 = 0.07	0.2937 0.4168 0.2099 0.0966 0.1751 **0.0260**

Tab. 4. Exemplified in matrix *D* the results of tests for linear acceleration of the mass a_2

Applied in the test parameters	The results $\cdot 10^{-3}$	Applied in the test parameters	The results $\cdot 10^{-3}$
backlash zone = 0.005, inertia moment J_1 = 0.7612, c = 0.8, epochs = 22, l = 0.9, δ_1 = 0.08	0.1350 0.0792 **0.0258** 0.1070 0.1759 0.1584	backlash zone = 0.005, inertia moment J_1 = 0.7612, c = 0.8, epochs = 25, l = 0.8, δ_1 = 0.08	0.1353 0.0794 **0.0258** 0.1072 0.1762 0.1588
backlash zone = 0.0075, inertia moment J_1 = 0.7612, c = 1.5, epochs = 22, l = 0.9, δ_1 = 0.08	0.2145 0.1587 0.1052 **0.0276** 0.0964 0.0790	backlash zone = 0.0075, inertia moment J_1 = 0.7612, c = 1.5, epochs = 27, l = 0.9, δ_1 = 0.07	0.2337 0.1729 0.1147 **0.0300** 0.1050 0.0860

4. CONCLUSIONS

The executed simulation tests of not established states of a complex model of an electromechanical system containing heavy nonlinearities have confirmed that using time-frequency methods with multistage decomposition of signal, or by a method of three-layer backpropagation neural network – these methods have been proved to be the effective research tools.

Using wavelet and neural network to analyze non-stationary signals of the tested electromechanical system enables effective limitation of possible losses resulting from consequences of increasing ranges of faults, which could lead to occurrence of a breakdown. Together with a simulation modeling this approach should be sufficient to enable complex verification process of the diagnostic algorithm.

Information contained in non-stationary diagnostic signals can be extracted from them both, through distributions of wavelet decomposition of coefficients as well as through values of weights of the neural network first layer for chosen parameters of a state, which describe physical quantities. This information may determine the type and place of occurrence of undesirable states of a complex nonlinear electromechanical system. Therefore, there is a possibility of obtaining correct diagnostic results concerning a system containing changing value of a viscous friction c for the given value of discontinuity in zero of the absorbing characteristics (c_1).

REFERENCES

Annamalai, B., & Swaminathan, S. T. (2016). Diagnostics of faults in induction motor via wavelet packet transform. *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, 01–06.

Chandralekha, R., & Yayanthi, D. (2016). Diagnosis of faults in three phase induction motor using Neuro Fuzzy Logic. *Journal of Applied Engineering Research*, *11*(8), 5735–5740.

Da Costa, C., Kashiwagi, M., & Mathias, M. H. (2015). Rotor failure detection of induction motors by wavelet and Fourier transform in non-stationary condition. *Case Studies in Mechanical Systems and Signal Processing*, *1*, 15–26. doi:10.1016/j.csmssp.2015.05.001

Douglas, H., Pillay, P., & Ziarani, A. (2003). Detection of broken rotor bars in induction motors using wavelet analysis. In *IEEE International Electric Machines and Drives Conference, 2003. IEMDC'03* (pp. 923–928). Madison, USA: IEEE. doi:10.1109/IEMDC.2003.1210345

Kowalski, Cz. (2005). Monitorowanie i diagnostyka uszkodzeń silników indukcyjnych z wykorzystaniem sieci neuronowych. *Prace Naukowe Instytutu Maszyn, Napędów i Pomiarów Elektrycznych Politechniki Wrocławskiej*, *57*(18), 226.

Orlowska-Kowalska, T., & Szabat, K. (2007). Neural-Network Application for Mechanical Variables Estimation of a Two-Mass Drive System. *IEEE Transactions on Industrial Electronics*, *54*(3), 1352–1364. doi:10.1109/TIE.2007.892637

Osowski, S. (1996). *Sieci neuronowe – w ujęciu algorytmicznym*. Warszawa: WNT.

Sridhar, S., Uma Rao, K., & Jade, S. (2016). Detection and classification of power quality disturbances in the supply to induction motor using wavelet transform and neural networks. *Balkan Journal of Electrical & Computer Engineering*, *4*(1), 37–44. doi:10.17694/bajece.62699

Zając, M. (2009). *Metody falkowe w monitoringu i diagnostyce układów elektromechanicznych*. Kraków: Wydawnictwo Politechniki Krakowskiej im. Tadeusza Kościuszki.

Applied Computer Science, vol. 14, no. 4, pp. 70–80
doi:10.23743/acs-2018-30

Submitted: 2018-11-30
Revised: 2018-12-07
Accepted: 2018-12-15

Maintenance department, Artificial neural network,
Manufacturing companies

Daniel HALIKOWSKI[*],
Justyna PATALAS-MALISZEWSKA [0000-0003-2439-2865][**],
Małgorzata SKRZESZEWSKA[**]

A MODEL FOR ASSESSING THE LEVEL OF AUTOMATION OF A MAINTENANCE DEPARTMENT USING ARTIFICIAL NEURAL NETWORK

Abstract

With regard to adapting enterprise to the Industry 4.0 concept, the first element should be the implementation and use of an information system within a manufacturing company. This article proposes a model, the use of which will allow the level of automation of a maintenance department to be forecast, depending on the effectivity of the use of the Manufacturing Executions System (MES) within a company. The model was built on the basis of the actual times of business processes completed which were supported by MES in the maintenance departments of two manufacturing enterprises using artificial neural network. As a result of research experiments, it was confirmed that the longer the time taken to complete business processes supported by MES, the higher is the degree of automation in a maintenance department.

[*] University of Applied Science in Nysa, Institute of Technical Science, ul. Armii Krajowej 7, 48-300 Nysa, Ponad, e-mail: daniel.halikowski@pwsz.nysa.pl
[**] University of Zielona Góra, Faculty of Mechanical Engineering, Institute of Computer Science and Production Management, Licealna 9 Street, 65-417 Zielona Góra, Poland,
e-mail: j.patalas@iizp.uz.zgora.pl, e-mail: m.skrzeszewska@wm.uz.zgora.pl

1. INTRODUCTION

Maintenance services are now an important element in the structure of a manufacturing company, since there is a need to adapt the operation of companies to the requirements of the Industry 4.0 concept. One of the elements of the "Industry 4.0" concept is the so-called "*smart factory*", which seeks to automate processes within an enterprise, particularly production processes which require special maintenance procedures which, in turn should be supported by an IT system. Within this context, developments, such as Industry 4.0 and information technology offer great potential (Seitz & Nyhuis, 2015).

The aim of the article is to build a model for assessing the degree of automation in a maintenance department, using artificial neural network (model AC-MD) based on an example using Manufacturing Executions Systems (MES). The implementation of MES in the maintenance department of a manufacturing company provides information that allows production operations to be optimised; at the same time, MES is used to support activities carried out in the maintenance department of production enterprises (Jacobson, Masson, Smith & Souza, 2005; Jacobson & Masson, 2006).

Good examples are to be found in the literature, on the use of artificial neural network in maintenance, for the monitoring of such as tool wear, diagnosing vibration in machining systems, in the thermal analysis of machines, in analysing other malfunctions affecting production, not to mention the geometrical analysis of the product as well as the diagnostics of the finished product (Gawlik & Kiełbus, 2012). The AC-MD model was built using artificial neural network on account of their usefulness both in reactive and preventive maintenance and, finally, in predictive maintenance.

Model AC-MD was formulated based on the real data, obtained from the maintenance departments of two, automotive industry companies. In the first manufacturing company analysed, workers at the strategic, tactical and operational levels in the maintenance department service 380 machines while in the second, the workers service 20 machines. For the purposes of the analysis, the time taken by the employees to perform the actions using MES, was adopted.

2. APPLICATION OF ARTIFICIAL NEURAL NETWORK IN MAINTENANCE

The automation of manufacturing processes, together with the minimisation of costs, (Wu, Tian & Chen, 2013) the shortening of production cycles and a reduction in the size of inventories, has improved the importance of repairs and maintenance as a service function of companies (Bojar & Żółtowski, 2011). This aspect of a company's operation is an important element in the development of a company in accordance with the Industry 4.0 concept. Ensuring continuity of the production process and the desired quality of manufactured products, requires the use of appropriate tools to supervise the condition of the machines and technical devices (Kosicka, Mazurkiewicz & Gola, 2016). The solution is the implementation of IT systems in maintenance departments, which support activities implemented at the operational, tactical and strategic levels. An example of an IT system that is used in a maintenance department is MES. MES systems enable the effective collection of data and information from business processes in real time; this data and information is collated in the production and maintenance departments and then subsequently transferred to other completed processes in the enterprise. Data and information on production can only be collected directly from machines and from employees working in production and maintenance departments.

The difficulty in applying the MES system is to assess whether the use of this system by employees is effective and whether it contributes to an increase in the degree of automation of a department. Does the application of MES allow the company to adapt to the requirements of Industry 4.0?

We are, therefore, looking for a model for assessing the degree of automation in a maintenance department, the use of which will allow an increase in efficiency in a company, when using the MES system.

Artificial neural network for the construction of the AC-MD model was selected. The operation of artificial neural network is based on the classification and generalisation of individual features or facts; their use is particularly justified in the case of the high complexity of a problem with little knowledge of the rules (Gawlik & Kiełbus, 2012). The use of neural network may refer, for example, to the problems of classifying or predicting time data for individual tasks. Examples of the use of artificial neural network for the parallel processing of data in real-time control systems (Wu, Tian & Chen, 2013; Li, Wang & Wang, 2017) indicate the validity of their use in the AC-MD model. In addition, examples of using artificial network for the purpose of identifying and forecasting the wear of machine elements or of the thermal distortion of grinders were found, as well as in the geometric identification of wear indicators, in such as the blades and the surface layer of an object. (Gawlik & Kiełbus, 2012; Wu, Tian & Chen 2013). With regard to the possibility of using artificial neural network in maintenance, additional examples of models were found, *viz.*, the analysis of the wear patterns of circular saw blades, forecasting the course of the abrasive wear of cutting tools' blades (Gawlik & Kiełbus, 2012), the grouping of machine construction elements (Lipski & Pizoń, 2014), infrared thermography for the detection of defects in elements (Huda & Taib, 2013) and the monitoring of the operation of ship machinery (Raptodimos & Lazakis, 2016).

3. MODEL AC-MD

Building an AC-MD model, using the artificial neural network, requires real input data to be obtained, regarding the realisation of selected business processes in the maintenance department of manufacturing enterprises. The following business processes were carried out in those departments whose execution times were defined over the course of one working week [in minutes] on three levels of management, namely, the strategic, tactical and operational levels. The activities designated were performed, either with the support of the MES system (*either partially or fully*), or without the support of the MES system, respectively at the three levels defined (Table 1):

Tab. 1. Activities carried out in the maintenance department of manufacturing companies

	Business processes	**Time taken up by activities over a week: Manufacturing company 1**			**Time taken up by activities over a week: Manufacturing company 2**		
		strategic level	**tactical level**	**operational level**	**strategic level**	**tactical level**	**operational level**
1.	Making entries in the records regarding the inspection of equipment/machines		100	100		25	25
2.	Making entries in the records regarding the testing/tuning of devices and machines		100	50		5	5
3.	Order management	100	30		100	50	
4.	Tracking the status of devices/machines in real time (on-line)	50	50	50	40	50	25
5.	Reporting the demand for external service	30	20		50	30	
6.	Monitoring/Tracking schedule/Production planning	10	20	25	10	20	
7.	Planning downtime	10	20		30	30	
8.	Identification of bottlenecks on each machine/device	30	60	100	40	50	60
9.	Registering parts/consumables for equipment/ machines		30	100		50	60
10.	Monitoring repairs to equipment/machines	50	20		55	50	40
11.	Review of technical documentation	30	50	50	35	50	30
12.	Checking the availability of parts in the warehouse	50	50	50	55	50	60

Tab. 1. Activities carried out in the maintenance department – continued

13.	Reporting the demand for parts/consumables	100	150	50	80	75	40
14.	Recording/making a selection from the list of actions performed	50	50	50	45	50	60
15.	Entering the write-up for devices/machines	25	50	50	25	50	30
16.	Recording the withdrawal of equipment/ machines from service	1	8		5	10	
17.	Reporting the availability for work of repaired devices/machines, *post* overhaul	100	60	50	60	50	70
18.	Simulation of re-tooling devices, machines/ production lines	60	8	120	120	150	150
19.	Generating reports for machines/devices	100	20		70	25	
20.	Signalling the downtime of equipment/machines	50	20	50	70	10	10
21.	Informing about failure/blockages	10	25	20	15	10	10
22.	Signalling/informing about the availability of equipment/machines/ production line	50	50	50	60	10	10
23.	Conducting on-line/video training	30	40		20	30	
24.	Training planning	3			5		
25.	Monitoring of training	8	5		10	10	
26.	Human resources planning	3	30		50	40	
27.	Creating procedures	15	20		20	30	
28.	Reporting/signalling improvements, such as modernisation, improvement of machines/devices	60	60	50	60	60	40
29.	Reporting/signalling solutions to improve work (e.g. information flow)	30	30	50	40	50	40
30.	Notification by SMS or e-mail about planned, preventive maintenance/repairs	10	10	6	15	10	20
31.	Generating the alarm manually on failure	10	10	12	8	15	20

Tab. 1. Activities carried out in the maintenance department – continued

32.	Generating the alarm automatically on failure	10	15	5	12	15	15
33.	Notification by SMS/e-mail of a failure	10	20	15	15	30	20
34.	Implementing improvement solutions (e.g. modernisation, improvement of machines, devices)	60	60	100	40	50	20
35.	Implementing solutions that improve work (e.g. information flow)	30	20		40	50	
36.	Monitoring the technical testing of equipment/machines	15	30	25	10	15	
37.	Running a repairs calendar	25	100		15	20	
38.	Access from the console to the desktop of another level	15	15		10	15	
39.	Monitoring MTTR indicator (Mean Time to Repair)	20			25	50	
40.	Monitoring MTTF indicator (Mean Time to Failure)	20			30	50	
41.	Monitoring MTBF indicator (Mean Time Between Failures)	20			30	50	
42.	Analysis of the availability of a device/ machine	20			20		
43.	Monitoring the OEE indicator (Overall Equipment Effectiveness)	15			60		
44.	Analysis of costs in a maintenance department	15			60		
45.	Recording accidents at work				30		
46.	Archiving data	100	100	25	180	120	10

To build the AC-MD model, an artificial, neural network with a linear, activation function was used. Weighting factors were determined, in the neuron training process, by supervised learning. To generate the neural network, **Matlab R2018a** was used with the built-in Neural Network Tool. The model of the neural network used in the study is shown in Fig. 1.

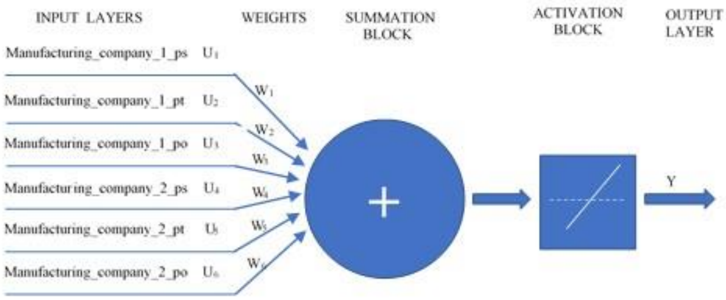


Fig. 1. Model of the neural network applied, based on a linear neuron (W1…W6 – weights; Y – level of automation of a maintenance department at the strategic, tactical and operational level in a company)

Based on the data analysed by the neural network with a linear activation function, the following AC-MD model was obtained.:

$$Y = 0,0294 \times P_1 + 0,1756 \times P_{17} + (-0,0201) \times P_{18} + (-0,2453) \times P_{20} + 0,0020 \times P_{34} + 0,1104 \times P_{46} \quad (1)$$

where: P_1 – business process: making entries in the records upon the inspection of equipment/machines,
P_{17} – business process: reporting the availability for work of repaired devices/machines, *post* overhaul,
P_{18} – business process: simulation of re–tooling devices, machines/ production lines,
P_{20} – business process: signalling equipment/machine downtime,
P_{34} – business process: implementing improvement solutions (e.g. modernisation, improvement of machines, devices),
P_{46} – business process: archiving data.

The AC-MD model includes business processes in a production enterprise, the implementation of which, influences the level of a maintenance department's automation index. It seems that these processes should be performed in an enterprise using the MES system and therefore their weekly time, performed with the support of the MES system should be increased. In order to verify the AC-MD model obtained, research experiments were carried out.

4. THE USE OF THE AC-MD MODEL

In order to find the answer to the question as to whether any increase in the use of the MES system in completing activities carried out in a maintenance department has any influence on the increase of the degree of automation of that maintenance department, the following working times were taken over the period of one working week [in minutes], Table 2.

Tab. 2. Time taken to complete activities carried out at manufacturing companies in a maintenance department – experimental data

	Business processes	Time taken up by activities over a week: Manufacturing company 1			Time taken up by activities over a week: Manufacturing company 2		
		strategic level	tactical level	operational level	strategic level	tactical level	operational level
E1	P_1 – Making entries in the records on the inspection of equipment/machines		100	100		25	25
E2			120	120		40	40
E3			140	140		60	60
E4			160	160		80	80
E1	P_{17} – Reporting the availability for work of repaired devices/machines, *post* overhaul	100	60	50	60	50	70
E2		120	80	70	80	70	90
E3		140	120	90	100	90	110
E4		160	140	120	120	110	130
E1	P_{18} – Simulation of re-tooling devices, machines/production lines	60	8	120	120	150	150
E2		80	20	140	140	170	170
E3		100	40	160	160	190	190
E4		120	40	180	180	210	210
E1	P_{20} – Signalling equipment/machine downtime.	50	20	50	70	10	10
E2		60	40	70	90	20	20
E3		80	40	90	110	40	40
E4		100	60	110	130	60	60
E1	P_{34} – Implementing improvements, such as, modernisation, improvement of machines, devices	60	60	100	40	50	20
E2		80	80	120	60	70	30
E3		100	100	140	80	90	50
E4		120	100	160	100	110	70
E1	P_{46} – Archiving data	100	100	25	180	120	10
E2		120	120	50	200	140	20
E3		140	140	70	210	160	30
E4		150	160	90	220	180	40

Based on the data (Table 2), the values forecast for the automation index of a maintenance department, were calculated (Figure 2).

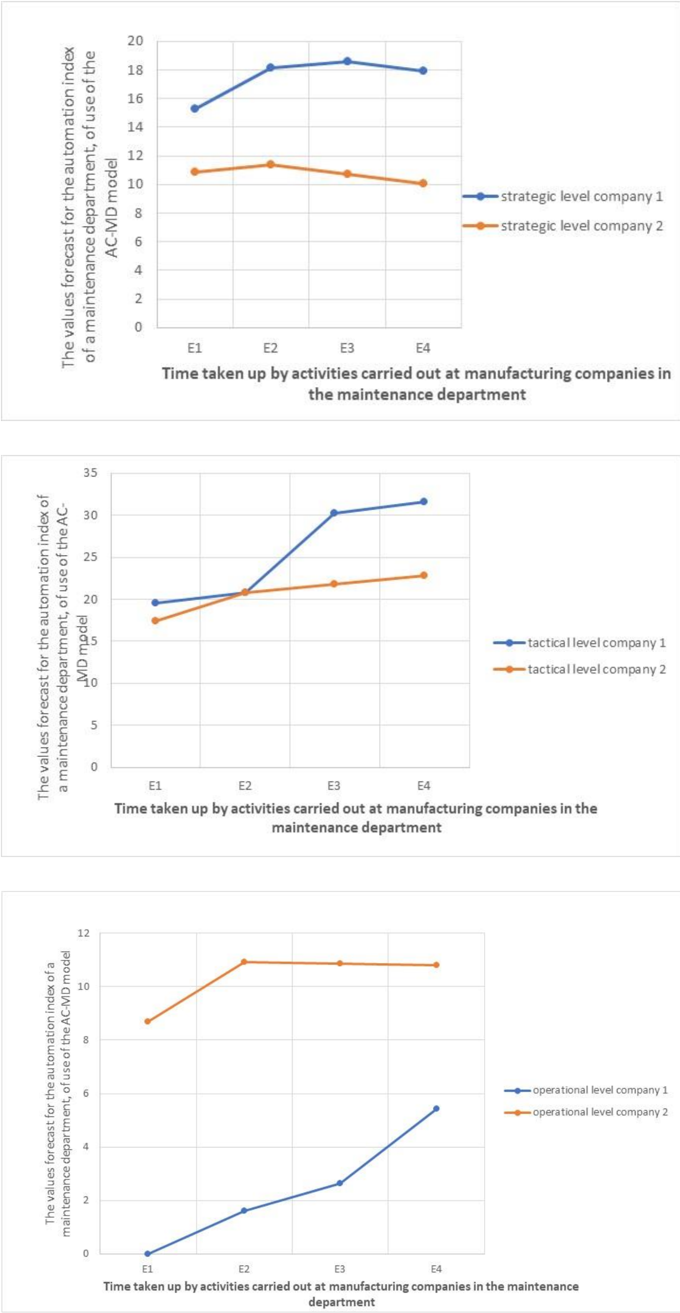


Fig. 2. The values forecast for the automation index of a maintenance department, at the strategic, tactical and operational level of use of the AC-MD model

Based on the data, it was found that the processes of making entries into records regarding the inspection of equipment and machines; the reporting of the availability for work of repaired devices and machines, *post* overhaul; the simulation of re-tooling devices, machines and production lines; the signalling of equipment and machine downtime; the implementation of improvements, such as modernisation and the improvement of machines, devices and archiving data, all have a significant impact on the level of the automation index in a maintenance department. In addition, the more these processes are supported by the MES system, the higher is the level of the automation of the whole department. Thus, increasing efficiency by implementing the use of MES, in a maintenance department, should be the first step in preparing an enterprise to implement the Industry 4.0 concept.

5. CONCLUSIONS

The article presents that the implementation and effective use of the **MES** system in a production company in a maintenance department can lead to an increase in the level of automation of the whole department. In further work, the authors will present IT tools and their application in the production enterprise which will allow the company to forecast the effects of investing in the implementation of the **Industry 4.0** concept.

REFERENCES

Bojar, W., & Żółtowski, M. (2011). Procesy wspomagania decyzji w zakresie utrzymania ruchu i eksploatacji maszyn. *Studia i Materialy Polskiego Stowarzyszenia Zarzadzania Wiedza, 40*, 71–84.

Gawlik, J., & Kiełbus, A. (2012). Zastosowania metod sztucznej inteligencji w nadzorowaniu urządzeń technologicznych i jakości wyrobów. In T. Sikora & M. Giemza (Eds.), *Praktyka zarządzania jakością w XXI wieku* (pp. 508-534). Kraków, Poland: Wydawnictwo Naukowe PTTŻ.

Huda, A. N., & Taib, S. (2013). Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment. *Applied Thermal Engineering, 61(2)*, 220–227. doi:10.1016/j.applthermaleng.2013.07.028

Jacobson, S., Masson, C., Smith, A. & Souza, J. (2005). AMR Research Report 18059, MES Market Rides Perfect Storm Through $1 B Barrier. *AMR Research*, 2–18.

Jacobson, S. & Masson, C. (2006). Eyelit: MES Lite: Building MES Composite Applications With Operations Process Management. Retrieved from http://eyelit.com/simon.html.

Kosicka, E., Mazurkiewicz, D., & Gola, A. (2016). Problemy wspomagania decyzji w systemach utrzymania ruchu. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska, 4*, 49–52. doi:10.5604/01.3001.0009.5189

Li, Z., Wang, Y., & Wang, K. S. (2017). Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario. *Advances in Manufacturing, 5(4)*, 377–387. doi:10.1007/s40436-017-0203-8

Lipski J., & Pizoń J. (2014), Sztuczna inteligencja w inżynierii produkcji. In J. Lipski, A. Świć, & A. Bojanowska (Eds.), *Innowacyjne metody w inżynierii produkcji* (pp. 11–24). Lublin, Poland: Wydawnictwo Politechniki Lubelskiej.

Raptodimos, Y., & Lazakis, I. (2016). An artificial neural network approach for predicting the performance of ship machinery equipment. In *Maritime Safety and Operations 2016 Conference Proceedings* (pp. 95–101). Glasgow, UK: University of Strathclyde Publishing.

Seitz K.-F. & Nyhuis P. (2015). Cyper-Physical Production Systems Combined with Logistic Models – A Learning Factory Concept for an Improved Production Planning and Control. *CIRP Procedia, 32*, 92–97. doi:10.1016/j.procir.2015.02.220

Wu, B., Tian, Z., & Chen, M. (2013). Condition-based maintenance optimization using neural network-based health condition prediction. *Quality and Reliability Engineering International,* 29(8), 1151–1163. doi:10.1002/qre.1466

Applied Computer Science, vol. 14, no. 4, pp. 81–92
doi:10.23743/acs-2018-31

Submitted: 2018-11-19
Revised: 2018-12-07
Accepted: 2018-12-10

Enterprise Resource Planning, SAP,
system implementation evaluation, improvement

Monika KULISZ [0000-0002-8111-2316]*

EVALUATION OF SAP SYSTEM IMPLEMENTATION IN AN ENTERPRISE OF THE AUTOMOTIVE INDUSTRY – CASE STUDY

Abstract

Modern businesses boost management by implementing integrated IT systems, such as highly popular in Poland – SAP software, to aid the enterprise resource planning (ERP). This paper evaluates the implementation of an ERP system at a steel wheel rims manufacturer and distributor. The main research method was questionnaire, conducted at the Logistics & Customer Service department. The data acquired from the conducted research were analysed and processed to evaluate the implemented solution. Areas for improvement were pinpointed and concerned the adjustment of the software solution to the needs of the enterprise.

1. INTRODUCTION

The development of integrated computer systems can be traced back to 1950s. The first representative was the MRP (Material Requirements Planning) system, which served as a tool for ensuring the supply and demand balance by scheduling the input material supply for manufacturing. The systems subsequently evolved to MRP II (Manufacturing Resource Planning) which additionally embraced the control over the resources of the enterprise, as well as production and distribution (Lambert, Calvasina, Bee & Woodworth, 2017). The 1980s was the time when ERP systems appeared for the first time, only to gain in popularity in the 1990s (Kumar, Maheshwari & Kumar, 2003).

* Lublin University of Technology, Faculty of Management, Department of Organisation of Enterprises, Nadbystrzycka 38, Lublin, Poland, e-mail: m.kulisz@pollub.pl

What is understood under the acronym ERP is the set of enterprise management tools for the real-time distribution, processing and integration of data. ERP solutions provide an exciting boost for the effectiveness of management through redesigning processes and creating a common environment for potential additional components, functions and applications, such as the internal communication tool. These systems may give the advantage point over the competitors in the market. They integrate supply processes, warehouse management, production, sales, marketing and distribution, logistics, as well as financial and human resources (Huang & Handfield, 2015). They may prove indispensible in aiding decision-making processes concerned with planning and managing the enterprise as they provide an easily accessible database of the organisation. ERP systems are implemented to integrate and consolidate the systems across an organisation into one that can meet and serve each department's unique needs and tasks. This centralised character of ERP IT solutions facilitates the operation of the entire enterprise, for instance by the common user interface (Olson, Johansson & De Carvalho, 2018). Furthermore, the ever-changing business market environment and technological evolution requires of such solutions that they are applicable on mobile devices, such as smartphones or tablets, and offer a desired degree of automation with a view to monitoring processes on the real time basis (Costin & Cojocaru, 2017).

In its infancy, the ERP system was introduced exclusively in large organisations, on account of high implementation costs. Before long the small and medium-sized enterprises niche was filled by the vendors of ERP systems, who have decided to accommodate the system to the emerging needs by simplifying and basing it on the WWW service. Free Open Source ERP solutions have then become feasible alternatives for such businesses. ERP systems are provided in a number of forms, such as: software as a service (SaaS), open-source software (OSS) and Service Oriented Architectures (SOA), as well as hybrids of these (Olson, Johansson & De Carvalho, 2018). In the SaaS model it is a cloud service provider that supplies the ERP system (Johansson & Ruivo, 2013). An important factor determining the effectiveness of SaaS model is its integration with other elements of the management system, which would be seriously limited but for the ERP system: scalability, efficiency, accessibility, configurability and resistance to damage. The wide range of available options appears to stress the importance of the solution vendor selection process (Kłosowski, 2013).

According to the 2017 Panorama Consulting Solutions Report on ERP Systems and Enterprise Software, the major motivation of companies behind the implementation of ERP systems is to improve business performance (17%), to ensure compliance (14%), to make employees jobs easier (14%) and to better integrate systems across locations (13%) (*Panorama Consulting Solutions*, 2017a).

Bearing in mind the substantial costs, the degree of complication and the time consumption that are inescapably connected with the implementation of the ERP systems, the selection of a suitable ERP solution vendor becomes crucial for the prosperity of an enterprise. (Kilic, Zaim & Delen, 2015). The global market of ERP systems is somewhat dominated by the giants like: SAP, Oracle, Microsoft (*Panorama Consulting Solutions*, 2017b). The proposed solutions are modular, which allows for adaptation of the system to the current needs of a customer. ERP implementation is furthermore a step towards standardisation of business processes within organisations and introduction of "better practices". Since virtually every enterprise has a specific modus operandi, the implementation of ERP solutions may take one of three forms: the processes of the enterprise are subordinated to the system, the ERP is adjusted to fit in the conditions of a given company, or by compromise. However, it ought to be remembered that adjusting the ERP system to an organisation ridden with bad practices may compromise the system's effectiveness and may eventually be counter-productive (Parthasarathy & Sharma, 2016).

Analysing ERP systems it becomes evident that these systems are highly desirable from the viewpoint of enterprise management, due to the fact that they: optimise the cycle time, the product development time, reduce the nuisance involved in performing routine operations, improve work quality, facilitate customer relation management, enable creating a unique database, which prevents data multiplication in the organisation and fosters integration of functions within the organisation (Kale, 2000). Despite these vast advantages, ERP systems are criticised on account of their high complexity which prevents the users from embracing the general view of the processes they are personally involved in, demands entering large amounts of data in order for the system to function properly and imposes to a greater or lesser extent systematic solutions on the users as "good practices" (Grabot, Mayere, Lauroua & Houe, 2014).

There is a wide range of methods that can be applied with the aim of conducting the analysis of available ERP systems, determining the criteria of selection and finally selecting the most suitable solution for a given company, including: a considerably popular exploratory interviews (Johansson & Ruivo, 2013) or community interview (Mishra & Mishra, 2011), or analytic network process (ANP) (Perçin, 2008), multi-criteria decision – AHP (analytic hierarchy process) (Kahraman, Beskese & Kaya, 2010), as well as hybrid methods combining the ANP multi-criteria decision methods – PROMETHEE (Kilic, Zaim & Delen, 2015).

Ruivo, Oliviera and Neto (2015) have conducted a comparative analysis of four existing systems: Microsoft NAV, SAP All-in-one, ORACLE JDE, and SAGE X3 in SMEs where they focused on such areas as: Collaboration, ERP use, Analytics, ERP value. The study showed that what appears to be the focal point for Dynamics and ORACLE is analytics system capability, whereas for SAP and SAGE – a greater collaboration of system capability. Furthermore, SAP and ORACLE appear to agree on the importance of greater ERP use, which is not the case for Dynamics and SAGE (Ruivo, Oliveira & Neto, 2015).

Once the decision is reached, the process proceeds to the next phase – the implementation. The implementation of ERP into a living enterprise tissue is as complex as it is time-consuming, not to mention the tremendous risk and substantial costs of introduction. The mean cost of ERP implementation reported by 342 respondents amounted to USD 1.3 million. The process proved equally time-inefficient, and the average from the reported answers was 16.9 months (*Panorama Consulting Solutions*, 2017a). The implementation of ERP is a process that essentially accounts for the specific character of the enterprise and its processes, engages vast resources, knowledge, skills and experience of the personnel. The process furthermore requires considerable capital investments, supply chain aptitude and transparency of the information circulating within (Huang & Handfield, 2015). Unfortunately, the success of such an endeavour is largely unpredictable, for which the significant part of responsibility may be attributed to the human factor (Jenko & Roblek, 2016). According to a last year's „Report on ERP Systems and Enterprise Software", 70% of respondents stated that they had achieved success, 26% - had conceded failure, and 4% were undecided; the numbers compared to the previous year have shown a 13% increase in success stories, but a 19% increase in the failed implementation category (*Panorama Consulting Solutions*, 2017a). From these results it becomes apparent that the success of ERP implementation cannot be guaranteed and involves a considerable risk factor.

In the body of literature in the field, there are numerous research works devoted to implementation of ERP and post-implementation verification. The latter however, are to a greater extent focused on measuring success based on success factors, such as Return on Investment or profit. Researchers employ diverse methods to determine and evaluate the project success in enterprises. Ngai (Ngai, Law & Wat, 2008) has introduced 80 critical success factors (CSFs) describing the implementation of ERP, which have provided the base for Sun, Wenbin and Rocky for the development of CSFs key performance indices (KPIs), which are associated with each stage of ERP implementation by ten local ERP experts. KPIs are determined with the Dumpster-Shafer method and evaluated by 10 experts. The models has been adjusted for the commercial use in a Chinese consulting agency. The solution is burdened with certain limitations in terms of appliance of the obtained results in other countries, owing to the fact that the relevance of certain factors may be culture- or country- specific (Sun, Ni & Lam, 2015). Huang and Handfield have endeavoured to assess the effect of ERP implementation and system vendor selection on the efficiency of the supply chain, based on the supply chain maturity model adapted from Gupta and Handfield (2011) (Huang & Handfield, 2015). Sudhaman and Thangavel have conducted analytical works into the efficiency of ERP projects, from the perspective of software quality, based on their quality measures (defect counts) by means of the Data Envelopment Analysis Constant Returns to Scale (DEA CRS) model and have identified the most efficient ERP projects (Sudhaman & Thangavel, 2015). The literature offers instances of ERP system users studies based on exploratory interview methods (Mahendrawathi, Zayin & Pamungkas, 2017; Gattiker & Goodhue, 2005), where along the technical elements the human factors is concerned. This is the method, which is presented in the following sections of this paper.

2. CASE STUDY: BACKGROUND

The object of the following study has been previously mentioned: this is a manufacturer of high-quality steel wheel rims for cars and LCVs. Currently the company is a supplier for a number of high profile automotive corporations, including: Fiat, Audi, Volkswagen, Maserati, Iveco, Renault, Hyundai, or Kia. In the organisational structure of the enterprise, we have the following departments: Maintenance, Wheel Rings and Assembly, Wheel Disk and Tooling, Painting and Packaging, Rim Shear Line, Logistics & Customer Service, Human Resources, Work Safety and Environment, Accounting and Finance, Purchasing, Sales, IT, Quality Assurance, Construction and Design Development, and New Investment Projects.

The reported study was carried out among the 30 employees of the L&CS (Logistics & Customer Service) department of the company. This department is primarily responsible for coordination of material and raw material flow in the supply chain. The employees plan transport and distribution of finished goods to buyers, and maintain relations with carrier partners. The employees of the department are in the constant search of new solutions for streamlining their work. Logistics Specialists strive to optimise transportation costs, and therefore closely cooperate with the Sales Department and Purchasing Department when designing new goods transporting routes and methods. The scope of their duties involves contacting clients and carriers. Production Planning Specialists analyse the supply chain data and are responsible for short-term and long-term planning of production, preparation of production schedules and control of production realisation and stock levels, as well as analysis of the workload of production lines with a view to their optimisation. All the listed activities aim to secure timely delivery and customer satisfaction.

The work at the company is organised around Microsoft Office, SAP ERP 6.0 and SAP Netweaver 7.0 software. The software is available under per seat license and connected in a local network. The software is in Polish language version. SAP ERP is obviously an ERP-class system for enterprise resource planning, and enterprise management support. The system holds and enables access to the company database, applications and various analysis tools, which significantly facilitates supply, production, customer care, sales financial and human resources management processes. By integrating the multiple separate modules the data is entered only once to provide access to all in need of accessing it. In addition the open architecture of this solution enables integrating the in-house system with external systems utilised by the company.

The SAP system implemented in the enterprise consists of the following modules: Corporate Services, Financials, Human Capital Management and Operations. SAP ERP Corporate Services is a module for the improvement of administrative processes regarding the company assets, project portfolio, business trips or foreign sales. Additional functions of the module include work safety and environmental protection compliance activities. SAE ERP Financials is a useful tool enabling in-depth analysis of financial data and finance control. The system automates the financial accounting and management and facilitates management of financial resources and settlement of accounts. SAE ERP Human Capital Management enables automation of all processes concerned with human resources management. SAE ERP Operations governs logistic processes and embraces the entire process: from placing the request to flexible invoicing and payment. The module optimises the material flow in the company and improves the management of designing and production processes and, above all, automates the production process (Missbach & Anderson, 2016).

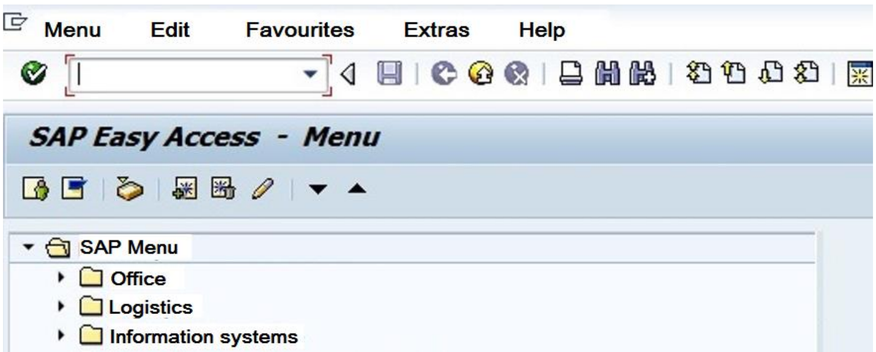


Fig. 1. Screenshot of the main dialogue box of the SAP system

The SAP system is navigated by transactions (commands), which are individual codes expressed by a series of letters and digits. Figure 1 shows the main dialogue screen appearing after logging into the system. The command field in the top left corner of the screen allows the user to input the appropriate code by hand. The main screen shows only those elements to which a given employee is permitted to access.

3. METHODOLOGY AND STUDY RESULTS

The conducted study was aimed to evaluate the implementation of the SAP system in an enterprise - manufacturer and distributor of steel wheel rims for the world leaders in the automotive industry. The assessment was based on the research questionnaire conducted among the personnel of the Logistics & Customer Service department.

The research methods employed in the study were: exploratory interview, observation and analysis of internal documentation obtained from the enterprise. The exploratory interview took the form of a questionnaire composed of 17 questions: 2 open-ended, 12 semi-open-ended, 3 close-ended, and 3 rating-scale questions. The initial question was the contingency question to determine the experience of employees with the ERP-type system. The respondents were asked about particular areas of their duties where the system was implemented, and the areas of information that it provides.

The survey questionnaire was presented to each of the 30 members of the L&CS department. Although this department has a young team, their experience with the system is quite extensive, as it is shown in Table 1. Based on the Table it becomes apparent that the selected employees are suitable for the revision of the system, as they are sufficiently experienced and because they depend on it on day-to-day basis they may share interesting insights into the errors of the system and suggest certain improvements.

Tab. 1. Years of work experience with the SAP ERP system

Years of work experience with SAP ERP	Number of replies	Percentage
< 2 years	9	30%
2–3 years	16	53.3%
> 3 years	5	16.7%

Subsequent questions asked the respondents to share their reflections regarding the intuitiveness of the system, its integration with other software already in use in the department (Microsoft Office), and to assess whether the implemented system has contributed to reducing working times of specific processes (Figure 2). Over half of the respondents considered SAP ERP as intuitive. One respondent developed the answer by adding that the interface was complicated. Admittedly, the system contains a number of elements in the dialogue boxes, which might be difficult to grasp for new users. Most commonly used basic options are intertwined with options for advanced users, which may prove particularly confusing in the interaction with the system. Over 75% of respondents regarded integration of SAP with Microsoft Office software as sufficient. The daily operations of the logistics department involves the Microsoft Office software either at the beginning or at the end of the business process conducted in the SAP environment. Initially, the data from Excel spreadsheets are automatically integrated into the SAP system, with no need for manual intermediation by generation of text files and importing duplicated files. The integration improves, inter alia, the effectiveness of report generation, without the need for mechanical copying of once entered data.

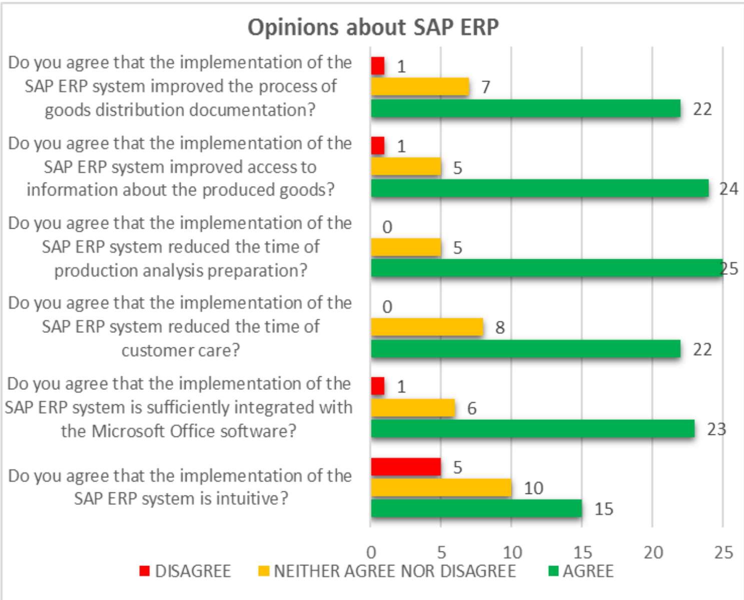


Fig. 2. Opinions about SAP ERP

The vast majority of the ones surveyed was of the opinion that the introduction of SAP ERP system had had a positive effect on reducing: the time of customer service (73.3%) and the time necessary to produce various analyses and reports (83.3%). A complete production report generated by the SAP ERP system considerably improves the daily work of production planners. Due to the versatility of this module, planners are presented with a tool enabling detailed analysis of data regarding particular production orders and generation of various cumulative reports, which may be further collated to highlight desired pieces of information. Further questions, shown in Figure 2, concerned the information provided by the SAP ERP system. In this case also the majority of respondents agreed that the access and flow of information regarding production (80%) and distribution (73.3%) have shown signs of improvement since the implementation of the system. SAP ERP enables quick and uncomplicated inspection of the correctness of production processes. The software may be used to control production on real-time basis, e.g. to assess whether a suitable remedial action should be commenced. The system enables timely reaction to off-norm situations, even when the process has already been started. If it were not for such a tool as SAP ERP, each necessary analysis would have to be manually conducted, which in the conditions of simultaneous manoeuvring several production lines, and daily production running into thousands, would have to involve extreme time waste, not to mention the lack of immediate reaction to unexpected scenarios. The SAP ERP system provides significant support to the entire organisation in terms of customer management. SAP provides data of sales profitability for particular customers, market segments or distribution channels. Therefore, it constitutes a basis for the optimisation of prices, e.g. by providing the knowledge of actual production costs.

The respondents of the survey were furthermore asked about the impact of SAP implementation on particular elements in different areas of the enterprise functioning (Figure 3). Similarly as in their response to previous questions, the vast majority of respondents have pointed to SAE ERP-induced improvements in material and raw material flow in the enterprise, control of stock level, due order delivery, customer management and closer returnable packaging control.

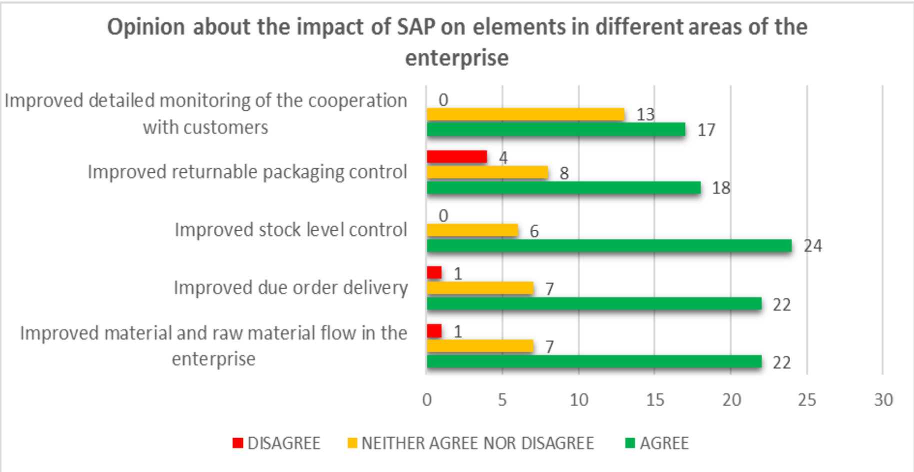


Fig. 3. Opinion about the impact of SAP on elements in different areas of the enterprise

SAP is a tool performing planning and scheduling of production, maintaining control over its profitability and controlling lots of semi-finished products and finished goods. The system aggregates data about the finished goods, such as the total number of pieces, according to storage, location, or other features of goods. The information about the storing location is immediately accessible, which has a positive effect on planning, coordination, and control of logistic processes involved in the goods flow. Customers regularly update their order schedules, which are instantly processed by SAP ERP to ensure that the current stock levels correspond to the customer order needs.

The key question asked in the survey required the respondents to reflect on the level of satisfaction with the SAP system. 70% of those who were interviewed confirmed that they were satisfied with the system, whereas 27% were neither for nor against the system, and one respondent expressed their disappointment with the work with the system. The results indicate that the system is positively perceived by the employees and that it fulfils its role in aiding the enterprise management. As this question allowed short commentary, some of the respondents shared their ideas on which areas of their work in the system could benefit from change. The major focus was predominantly ascribed to the interface, which was described as outdated and overloaded with information, containing excessive amount of transactions (commands), which furthermore frequently multiply the information.

4. STUDY CONCLUSIONS

The results obtained from the exploratory interview, the observations and analysis of documentation obtained from the enterprise lead to the formulation of several conclusions. SAP system is an indispensible tool, which organises the day-today work activities of the enterprise. This may be inferred from the answers of the respondents, the vast majority of whom expressed their positive opinion of the system. ERP software is attracting an increased amount of attention from production/distribution companies, which consider it as an indispensible tool at work. Characterised by high complexity and tremendous production or distribution data storage capacity, the SAP ERP system is equipped in data filtering and configuration solutions, by means of transactions (commands), which significantly shorten the analysis preparation times. The considerable variety of transactions allows formulation of high-complexity analyses and purpose-specific datasets, which furthermore are the proof of high flexibility of the system.

Apart from the indisputable advantages of the system, its interface is regarded as outdated and overloaded with information. The former causes problems remembering the transactions, which could be resolved by a greater flexibility of the system in editing and customising the interface (i.e. the list of transactions) as per particular user needs. From the technical perspective, the change could consist in creating a "Most used" folder in the main menu, which would hold the list of transactions of greatest relevance to a particular employee, his position and expertise. This solution would prevent errors in entering transaction codes and shorten the time required to find the desired command. The reduced list of available options would enable instant access to key transactions, whereas the user would not be required to memorise the transaction codes. Moreover, the outdated interface indicates the need to implement a systematic update that would elevate the interface to current visual standards.

Furthermore, the SAP ERP system could become more accessible if the mobile version of the application were introduced and installed on company smartphones. Such apps increase the flexibility of employees and ensure a significantly faster reaction to unexpected failures, which are not uncommon in the analysed industry. The mobile application would allow the employees/users to access the system data at any given time, including off work. In critical situations a member of the logistics team would not have to appear at the company to e.g. correct the documentation, and would be able to perform any actions remotely from home. This solution could prove particularly useful in emergency, which is an everyday occurrence in transport and logistics.

The drawbacks of the system discovered in the process, are far from outweighing the benefits of the solution, such as reduction in the data processing times, in reporting and production analysis. The logistics team members could devote the time saved by the automation of SAP ERP system to other burning issues in other areas of company operation, rather than to multiplying the existing data entries in spreadsheets or files.

The results from this study indicate that ERP systems in modern enterprises may prove irreplaceable and are crucial to their smooth daily operation. The wide array of potential applications and modules enables adjusting the system to various enterprises and their specific needs. Together with relieving employees of certain daily duties, the system is an effective tool in shortening the time necessary to create various production analyses. The system provides support to workers and the source of various information. Over 70% of the respondents claim that their work would be virtually impossible without the aid of the ERP system. This indicates that the system is highly useful and is regarded as an indispensible element of everyday work.

REFERENCES

Costin, B., & Cojocaru, D. (2017). Integration of metrology applications in the calibration reservoir suites using SAP Fiori, portal and cloud. A study case. *2017 21st International Conference on System Theory, Control and Computing (ICSTCC)* (pp. 297–302). Sinaia: IEEE. doi:10.1109/ICSTCC.2017.8107050

Gattiker, T., & Goodhue, D. (2005). What Happens after ERP Implementation: Understanding the Impact of Interdependence and Differentiation on Plant-Level Outcomes. *MIS Quarterly, 29*(3), 559–585. doi:10.2307/25148695

Grabot, B., Mayere, A., Lauroua, F., & Houe, R. (2014). ERP 2.0, what for and how? *Computers in Industry, 65*(6), 976-1000. doi:10.1016/j.compind.2014.02.017

Huang, Y. Y., & Handfield, R. (2015). Measuring the benefits of ERP on supply management maturity model: a “big data” method. *International Journal of Operations & Production Management, 35*(1), 2–25. doi:10.1108/IJOPM-07-2013-0341

Jenko, A., & Roblek, M. (2016). A Primary Human Critical Success Factors Model for the ERP System Implementation. *Organizacija, 49*(3), 145–160. doi:10.1515/orga-2016-0014

Johansson, B., & Ruivo, P. (2013). Exploring Factors for Adopting ERP as SaaS. *Procedia Technology, 9,* 94–99. doi:10.1016/j.protcy.2013.12.010

Kahraman, C., Beskese, A., & Kaya, I. (2010). Selection among ERP outsourcing alternatives using a fuzzy multi-criteria decision making methodology. *International Journal of Production Research, 48*(2), 547-566. doi:10.1080/00207540903175095

Kale, V. (2000). *Implementing SAP™ R/3: The Guide for Business and Technology Managers* (1 ed.). Indianapolis: Sams Publishing.

Kilic, H. S., Zaim, S., & Delen, D. (2015). Selecting "The Best" ERP system for SMEs using a combination of ANP and PROMETHEE methods. *Expert Systems with Applications, 42*(5), 2343–2352. doi:10.1016/j.eswa.2014.10.034

Kłosowski, G. (2013). Cloud Manufacturing Concept as a Tool of Multimodal Manufacturing Systems Integration. *Foundations of Management, 4*(1), 17–42. doi:10.2478/fman-2013-0002

Kumar, V., Maheshwari, B., & Kumar, U. (2003). An investigation of critical management issues in ERP implementation: emperical evidence from Canadian organizations. *Technovation, 23*(10), 793–807. doi:10.1016/S0166-4972(02)00015-9

Lambert, S., Calvasina, R., Bee, S., & Woodworth, D. (2017). Assembly FG: An Educational Case on MRP II Integrated within ERP. *Accounting Perspectives, 16*(1), 43–46.

Mahendrawathi, E., Zayin, S., & Pamungkas, F. (2017). ERP Post Implementation Review with Process Mining: A Case of Procurement Process. *Procedia Computer Science, 124*, 216–223. doi:10.1016/j.procs.2017.12.149

Mishra, A., & Mishra, D. (2011). ERP Project Implementation: Evidence from the Oil and Gas Sector. *Acta Polytechnica Hungarica, 8*(4), 55–74.

Missbach, M., & Anderson, G. (2016). *SAP w 24 godziny.* Gliwice: Helion.

Ngai, E., Law, C., & Wat, F. (2008). Examining the critical success factors in the adoption of enterprise resource planning. *Computers in Industry, 59*(6), 548–564. doi:10.1016/j.compind.2007.12.001

Olson, D., Johansson, B., & De Carvalho, R. (2018). Open source ERP business model framework. *Robotics and Computer-Integrated Manufacturing, 50,* 30–36. doi:10.1016/j.rcim.2015.09.007

Panorama Consulting Solutions. (2017a). *Report on ERP Systems & Enterprise Software.* Greenwood Village: Panorama Consulting Solutions.

Panorama Consulting Solutions. (2017b). *Top 10 ERP Systems Rankings Report.* Greenwood Village: Panorama Consulting Solutions.

Parthasarathy, S., & Sharma, S. (2016). Efficiency analysis of ERP packages – A customization perspective. *Computers in Industry, 82,* 19–27. doi:10.1016/j.compind.2016.05.004

Perçin, S. (2008). Using the ANP approach in selecting and benchmarking ERP systems. *Benchmarking, 15*(5), 630–649. doi:10.1108/14635770810903196

Ruivo, P., Oliveira, T., & Neto, M. (2015). Using resource-based view theory to assess the value of ERP commercial-packages in SMEs. *Computers in Industry, 73*, 105–116. doi:10.1016/j.compind.2015.06.001

Sudhaman, P., & Thangavel, C. (2015). Efficiency analysis of ERP projects–software quality perspective. *International Journal of Project Management, 33*(4), 961–970. doi:10.1016/j.ijproman.2014.10.011

Sun, H., Ni, W., & Lam, R. (2015). A step-by-step performance assessment and improvement method for ERP implementation: Action case studies in Chinese companies. *Computers in Industry, 68,* 40–52. doi:10.1016/j.compind.2014.12.005

Applied Computer Science, vol. 14, no. 4, pp. 93–108
doi:10.23743/acs-2018-32

Submitted: 2018-09-06
Revised: 2018-12-01
Accepted: 2018-12-16

Business standards, Interoperability, Canonical Data Models, Graphs, Graph Databases, Graph Transformations

Leszek JASKIERNY [0000-0002-9451-8569]*

REVIEW OF THE DATA MODELING STANDARDS AND DATA MODEL TRANSFORMATION TECHNIQUES

Abstract

Manual data transformations that result in high error rates are a big problem in complex integration and data warehouse projects, resulting in poor quality of data and delays in deployment to production. Automation of data transformations can be easily verified by humans; the ability to learn from past decisions allows the creation of metadata that can be leveraged in future mappings. Significant improvement of the quality of data transformations can be achieved, when at least one of the models used in transformation is already analyzed and understood. Over recent decades, particular industries have defined data models that are widely adopted in commercial and open source solutions. Those models (often industry standards, accepted by ISO or other organizations) can be leveraged to increase reuse in integration projects resulting in a) lower project costs and b) faster delivery to production. The goal of this article is to provide a comprehensive review of the practical applications of standardization of data formats. Using use cases from the Financial Services Industry as examples, the author tries to identify the motivations and common elements of particular data formats, and how they can be leveraged in order to automate process of data transformations between the models.

[*] AGH University of Science and Technology, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Department of Computer Science,
e-mail: leszekj@agh.edu.pl

1. INTRODUCTION

Although from a business perspective data models are created to support particular business objectives, from the practical perspective they must be focused on particular outcomes and selected groups of the end users. Both, technical complexity of the model and level of details included in the data model are dependent on the reason why the model was created and on its target audience. Looking from this perspective, models can be divided into sub-categories such as: conceptual models, logical models, and physical models. From a technical perspective, models can be defined as structured models, requiring strict definition of the details of the model, and non-structured (conceptual) models, built from more generic, loosely coupled objects. Conceptual models are created using various modeling techniques that don't enforce formal validation of the model. Structured logical models are created using modeling techniques supporting formal validation of the model. A good example of such a model is a UML class diagram

Finally, physical models can be represented by various schema definitions, e.g. XML Schema Definition (XSD), that specify how to formally describe the elements in an Extensible Markup Language (XML) document.

Such a formal definition is mandatory for documents being exchanged between applications. Further specifications, such as Web Service Description Language (WSDL), are built on top of XSD, to specify functional (business related) and non-functional (IT systems related) aspects of the data exchanged between applications.

2. REVIEW OF DATA FORMAT STANDARDIZATION METHODS AND TECHNIQUES

The following chapter provides an overview of data modeling languages and standards.

2.1. Data Format Description Language

Data Format Description Language (DFDL), published as an Open Grid Forum Proposed Recommendation in January 2011, is a modeling language for describing general text and binary data in a standard way. A DFDL model (or schema) allows any text or binary data to be read (or "parsed") from its native format and to be presented as an instance of an information set.

Unlike XSD, which is usually created to model an existing business environment, DFDL takes a practical approach, building a model based on the existing data structures where, in some cases, source data might not be well formatted.

An information set is a logical representation of the data contents, independent of the physical format. For example, two records could be in different formats, because one has fixed-length fields and the other uses delimiters, even though they contain exactly the same data and are both represented by the same information set.

2.2. Dataset Structure Definition

Dataset Structure Definition (DSD) describes how information in a specific dataset is structured. Knowledge of the structure is important, because it allows desired information to be filtered out and very precisely limited specific dimensions based on selected criterions. In addition to DFDL, DSD provides a mechanism to model, not only particular data items, but also complex data structures.

2.3. UML

Unified Modeling Language is a general-purpose modeling language, intended to help visualize the design of a system. Data models, as part of the overall design of a system, can be visualized using Class Models. Class Models significantly differ from business conceptual Entity-Relationship Diagrams because of their focus on object-oriented design. Class models can be used at any stage of the modeling process, i.e. to create conceptual models, logical models and physical models. Although Class Models have several drawbacks they set a foundation for other notations, e.g. XSD.

2.4. XML

Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable (Thompson & Lilley, 2014). The design goals of XML emphasize simplicity, generality, and usability. It is a textual data format with strong support via Unicode for different human languages (Bray, Paoli, Sperberg-McQueen, Maler & Yergeau, 2008).

2.5. XML Schema Definition (XSD)

XSD, a recommendation of the World Wide Web Consortium (W3C), specifies how to formally describe the elements in an Extensible Markup Language (XML) document. Such a formal definition is mandatory for the documents being exchanged between applications. Further specifications, such as Web Service Description Language (WSDL), are built on top of XSD to specify functional (business related) and non-functional (IT systems related) aspects of the data exchanged between applications.

3. INDUSTRIAL DATA MODELS

The following chapter provides an overview of the standards from the Financial Services Industry, influenced mainly by the commercial organizations, such as Society for Worldwide Interbank Financial Telecommunication (SWIFT), who provides a network that enables financial institutions worldwide to send and receive information about financial transactions in a secure, standardized and reliable environment. Over the past several years, SWIFT provided strong commercial support, leading to the creation of widely adopted messaging standards, e.g. ISO 15022 and ISO 20022. The following examples aim to provide an overview of those standards, in order to identify common characteristics driving particular vendors and particular industry solutions.

3.1. SWIFT MT

SWIFT allows financial and non-financial institutions to transfer financial transactions through a financial message ("SWIFT", 2018).

While SWIFT started primarily working with the simple payment instructions, it now sends messages for wide variety actions including security transactions and treasury transactions. Nearly 50 percent of SWIFT traffic is still for payment-based messages, 43 percent now concern security transactions, and the remaining traffic flows to treasury transactions.

SWIFT can be considered as a true pioneer of the global standardization of data formats on the industry level. All the messages are built to comply with various needs and regulations, while maintaining core compatibility across the industry.

The main benefit of SWIFT messages is their business content, while standardization of the data format is important, the main challenge is in ensuring that message's business content meets the following criterion:

- Universal core business entities are shared by the whole industry,
- Extendable, specific extensions are built on top of the core components of the model in order to preserve compatibility, while maintaining customer-specific features,
- An open model has to be prepared for future extensions and changes, while maintaining its core compatibility.

By setting-up the above foundations, SWIFT made an important step towards industrialization of the data exchange processes. SWIFT messages consist of five blocks of data including three headers, message content, and a trailer. Message types are crucial to identifying content. All SWIFT messages include the literal "MT" (Message Type). This is followed by a three-digit number that denotes the message category, group and type.

3.2. ISO 20022

CIOs and IT architecture experts from leading banks and other major financial institutions, along with independent software vendors, systems integrators, and SAP, have joined in a collaborative effort to shape the future of enterprise service-oriented architectures (enterprise SOA) in the banking industry. This joint effort – known as the SAP industry value network (IVN) for banks – was launched in September 2005 and had 35 members ("ISO 20022", 2018). Although the main focus of IVN was creation of the reusable services, creation of the appropriate data model was necessary in order to assure data consistency and interoperability of those services. SWIFT, as one of the strong contributors to IVN for Banking, took over work on the definition of a comprehensive data model that would cover main activities of the banks. This effort greatly enhanced the previous standardization efforts of SWIFT, mainly related to payments and asset transfers. The ISO 20022 standard is described in the document "ISO 20022 Financial Services – Universal financial industry message scheme".

3.3. OASIS Universal Business Language (UBL)

UBL, the Universal Business Language, defines a royalty-free library of standard XML business documents supporting digitization of the commercial and logistical processes for domestic and international supply chains such as procurement, purchasing, transport, logistics, intermodal freight management, and other supply chain management functions (Holman, 2018). UBL can be thought of as a language that allows disparate business applications and trading communities to exchange information along their supply chains using a common format. UBL also provides the opportunity to end the debate over standards for business document formats that has discouraged the adoption of new technologies for conducting business in the digital age.

4. PROPRIETARY DATA MODELS

Despite high level of standardization, particular organizations (banks, insurance companies, payment providers, etc.) still tend to use proprietary data formats that meet particular business requirements and allow them to distinguish from their competitors.

Because standardization processes are expensive, there is always a need for strong commercial support in order to produce tangible results. Commercial institutions are funding developments of internal data models, that are focusing on specific business requirements and following the particular technical limitations of a given organization (McKnight, 2014). Those models (such as ISO20022) are often used as a central model, spanning other data models used in the organization (e.g. models defined by particular software systems) and can be referenced as Canonical Data Models (CMD) (Roman, 2006).

Proper construction of the CMD is especially important from the perspective of ensuring efficient and flawless Data Transformation Processes. A systematic approach to CMD creation, resulting in the highly optimized data model, can be done using the following steps:

- Review the Domain Models already used in the Organization,
- Prepare a common data dictionary, including the main business and technical datatypes,
- Define business areas and main groups of data, together with a dependency matrix between those areas,
- Review other relevant industry models, e.g. SWIFT (ISO 15022 and ISO 20022),
- Map business requirements on the CMD templates created during past projects,
- Build versioning mechanisms directly into the model (CMD),
- Extend the meta-data related to the CMD, e.g. define defaults for particular datatypes, iterate and enhance the model with business owners.

Some of the commercial models have been wrapped up and packaged in the form of the templates that are distributed on a commercial basis. An example is ADRM Software – the leading independent provider of large-scale industry-specific data models. ADRM Enterprise Data Models provide the reference for related industry business areas, data warehouses and data mart models ("ADRM Software", 2018). The ADRM Enterprise Data Model incorporates integrated data requirements taken from the best-practices developed by organizations from particular industries, combined within a single model consisting of approximately 400 entities and 2,500 attributes. It is the essential data model for strategic planning, communicating information requirements across the organization, developing integrated systems and organizing data in the Business Area, Data Warehouse and Data Mart models. Each Enterprise Data Model is built upon a common core of entity building blocks, which contributes those same common entities for the construction of business area, data warehouse, data and application models.

A common set of core entities enables the related models to be consistent and extendable across the industries. The same common core of entities is used wherever applicable in other industries, thus providing a means of integrating data across different industries or lines of business. Such a data model has several common design characteristics:

- Industry-specific design,
- Comprehensive business area coverage,
- Is fully-attributed,
- Provides complete and detailed definitions,
- Is semantically clear and easy-to-understand,

– Reflects current industry data best-practices,
– Provides flexible and extendible design,
– Utilizes industry-standard data whenever possible,
– Can be presented using large format graphic representation,
– Supports a wide audience of interests,
– Is integrated with Business Area, Data Warehouse and Data Mart models.

5. BANK-MF PROPRIETARY MESSAGE FORMAT OF BANK X

Bank-MF (or Bank Message Format) is used as an example of a proprietary data format; the following example is not directly related to any particular implementation, but rather summarizes common elements of various implementations of proprietary data models.

Note: Data Model and Message Format are used here as synonyms; this is because focus of this article is on the data models used for the exchange of information between various IT systems. Usually this kind of data exchange is done using messages built upon technical formats (such as XML or JSON) and business data models.

In order to understand the complexity of the model, let's analyze the following diagram that presents the components of the sample solution: The Lego bricks represent messages compatible with Bank-MF data model (or Canonical Model), while wooden bricks and Post-It notes represent proprietary data formats, derived from particular IT systems.

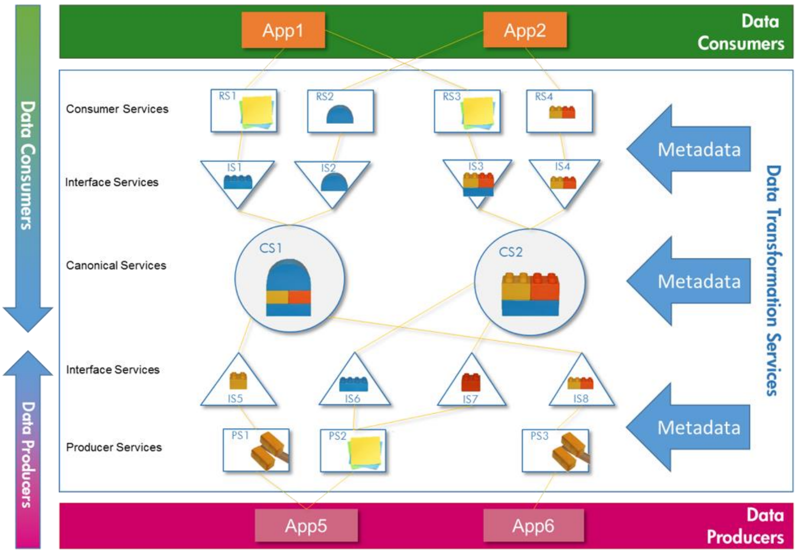


Fig. 1. Data transformation via Canonical Model

Although particular proprietary models are representing the same business data, there is a need for additional knowledge (referenced here as Metadata), in order to translate data between particular formats. The most important observation is that a canonical model is central to all other models, i.e. trans-formation between any pair of the proprietary models always goes via a Canonical Data Model. In order to facilitate mediation between a large number of proprietary models, Bank-MF has to meet the same criteria, as defined in a foundation of the SWIFT MT messages: universality, extensibility and openness.

The sample model referenced in this article was defined as a complex XSD schema, with clearly defined common components, basic data types, complex data types and orchestration of the messages into complex structures, representing particular business objects.

6. CHARACTERISTICS OF THE GRAPH MODELS

Graph theory is the study of graphs, which are mathematical structures used to model pairwise relations between objects. A graph in this context is made up of vertices, nodes, or points, which are connected by edges. A graph may be un-directed, meaning that there is no distinction between the two vertices associated with each edge, or its edges may be directed from one vertex to another (Angles & Gutierrez, 2008). The edges may be directed or un-directed.

In case of this article, we will focus on directed graphs, defining a clear relation between a parent node and a child node. In addition, each node of a graph will be equipped with a set of the attributes, defining proprietary characteristics of this node.

7. TRANSFORMATION OF THE DATA MODEL TO ITS GRAPH REPRESENTATION

As described earlier, there is large number of various data modeling standards and patterns common in multiple commercial implementations. In order to enable automated transformation of the data model, first we need to perform initial transformation between the actual definition of the model and reference definition, expressed in form of the graph. Alongside this transformation, we are also trying to understand the business context of given data model. Understanding the context is mandatory, because proper transformation has to be based on the business context of the data, not only on its technical implementation (Sleger, 2010).

In this article, we will attempt to provide a generic prescription for how to map any proprietary data model on the formal definition of the graph. The following chapter describes formal methods and techniques used to transform structured (e.g. XSD), or unstructured (e.g. NoSQL) data models into graph representations.

Further on, this process will be called universalization (or initial transformation) of the data model, i.e. transformation of a given model to its universal form. Normalized models would be compatible in terms of formal definitions, such as definition of the dependencies, list of attributes, etc.

Unlike the highly automated Data Model Transformation Process described later in this article, Initial Transformation might require signification support from the Subject Matter Expert (SME) knowledgeable about the business context of the model being transformed. The ultimate goal is to make the transformation process automated to its maximum possible extent. In order to achieve a high level of automation, we need to build a knowledge base containing already processed patterns and templates.

Presented below is an example from the financial industry.

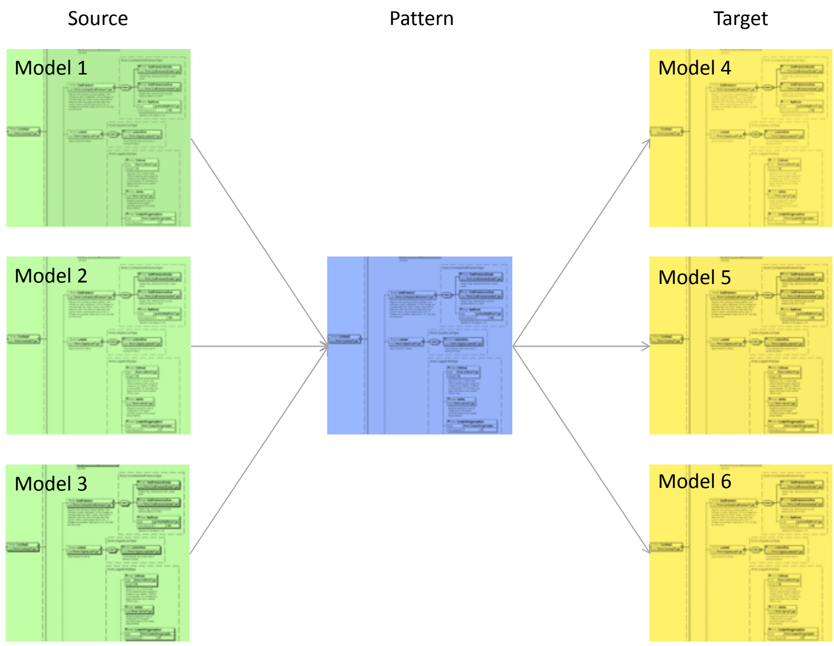


Fig. 2. Building a pattern from linked data models

In this example, the type of the data model can be recognized by detecting the dependencies between data elements in the model and by detecting the element names. The following list of dependencies: Contract_Acc_Currency can be matched against patterns and can by qualified as a financial data model.

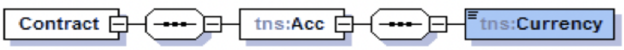


Fig. 3. Example of data dependencies

Knowing that, we can significantly limit the number of patterns that will be used for further analysis of the model. Following on that, we can use a limited dictionary to match the names of the attributes with the patterns. Once this simple model is analyzed and confirmed by the SME, it will be added to the knowledge base of patterns. In the future, similar patterns will be recognized automatically.

Initial transformation is a multi-step process, designed to extract important business information from the source model (S) and make sure that none of the business information is lost. At the same time, non-business information (e.g. headers, routing information, signatures, etc.) might be removed from the source model before its normalization.

The following graphic illustrates the process described below:

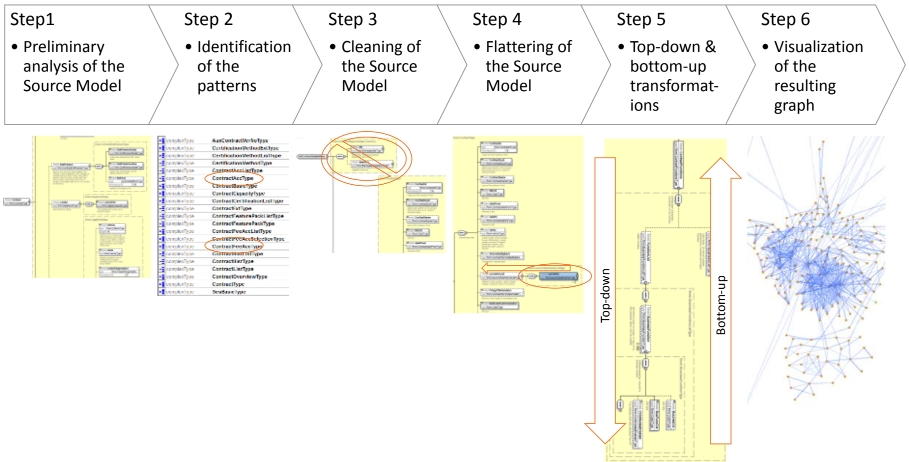


Fig. 4. Transformation of the Data Model into Graph representation

Step 1. The first step leads to the identification of the major artifacts; this step can be performed by simple counting of the attributes used in the model. E.g. a large number of occurrences of the complex data type called 'account' suggests that this might be one of the key attributes of the model. Following on this process, we can build the list of attributes that are most popular in the source model.

Step 2. During the second step of universalization, we are looking for given patterns and pre-defined elements (see the example earlier in this paragraph). The goal of this process is to associate initial transformation with the right set of dictionaries and patterns. In other words, we need to understand the overall business context of the data model in order to apply the right set of the transformations.

Step 3. The third step of data model universalization is focusing on cleaning-out unnecessary elements from the model, especially technical artifacts. The importance of this step would differ significantly between various models. Technical (non-functional, non-business) elements of the data model could be either entirely removed, or replaced with standard place-holders, in order to simplify further processing while not losing any important artifacts. This stage of the process might require manual input from the expert.

Step 4. The next step is about 'flattening' of the model. Highly normalized data models tend to be extremely complex and hard to visualize. Simplification focuses on the removal of unnecessary groupings, while maintaining the original business meaning of the model.

Step 5. Now comes an important question: where to start the 'universalization' process? Two possible choices are: top-down or bottom-up. The first option focuses on the overall meaning of the model, where each node is further enriched by the definition of its child elements. The bottom-up approach focuses on the details that could be further composed into the meaningful entities. Although the top-down option is preferred, certain element of the bottom-up approach is necessary in order to preserve the right amount of detail. Graph notation is a very good way to explain the dependencies within the model, and this is why the top-down approach seems to be a better choice.

Step 6. The final stage of universalization allows us to generate a consistent graph-view of the source model. Using a graph-view, we can easily understand the overall complexity of the model, its bottlenecks, cross-dependencies and the overall structure.

7.1. Preparation for the transformation of the data model

The following example presents the transformation of the physical data model (that could be represented by Data Definition Language) to its graph equivalent. Transformation between the flat structure and graph have been performed using metadata generated from the Reference Data Model.

The following changes have been applied as a part of initial transformation of the model:

1. Element names have been changed, according to the data stored in the dictionary.
2. Basic data structures have been created from the flat structure, based on the metadata stored in the dictionary.

Customer	
First_name	Str(20)
Middle_Initial	Str(1)
Last_name	Str(30)
Cust_ID	Str(14)
Street	Str(35)
House	Str(6)
State	Str(2)
City	Str(30)
Cust_type	Int

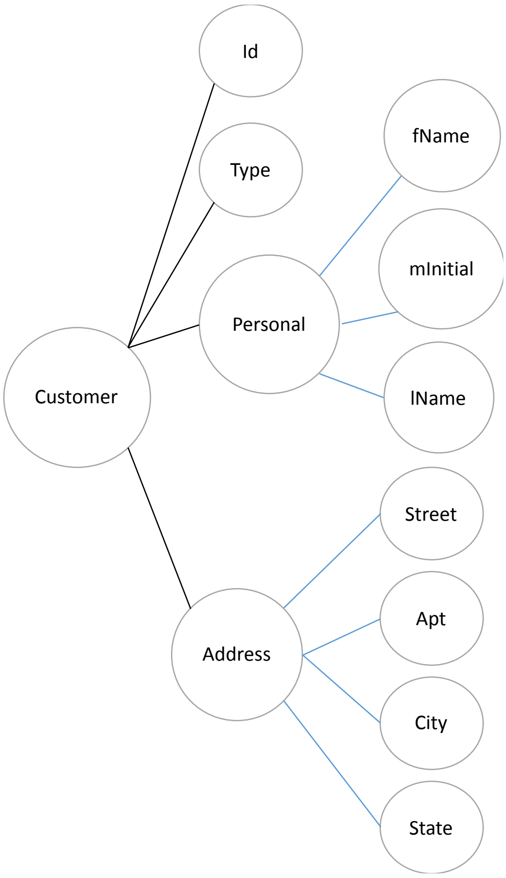


Fig. 5. Visualization of the dependencies in the data model

The resulting graph doesn't represent Reference Data Model yet, but is using the naming convention and basic structures, compatible with the template. Further work would be required to perform a full transformation of the source model to the Reference Data Model.

7.2. Data Model Transformation Processes

Transformation means that two Data Models, that are compatible on a business level, can be equipped with a set of the rules that enable transformation between those models (Kotulski, 2013). Those two models are further referred to as a Source Model (S) and a Target Model (T). Transformation between these models can be expressed as:

$$S \rightarrow T$$

The assumption is that both models S and T, are already “normalized” to the graph-representation.

In order to prove that the models are compatible (i.e., can be transformed), we need to analyze each graph, looking for the productions (results of graph’s transformations) that have to be used to generate this graph. This set of productions, for the source model, can be defined as:

$$SL \rightarrow SR$$

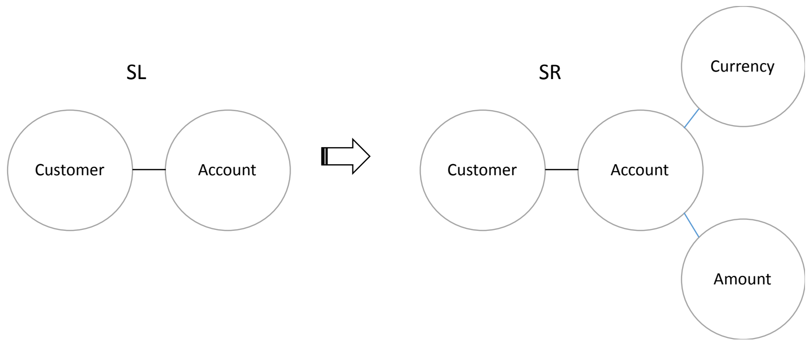


Fig. 6. Graph production on the source model

Similar analysis would be performed on the target model. The set of productions for the target model can be defined as:

$$TL \rightarrow TR$$

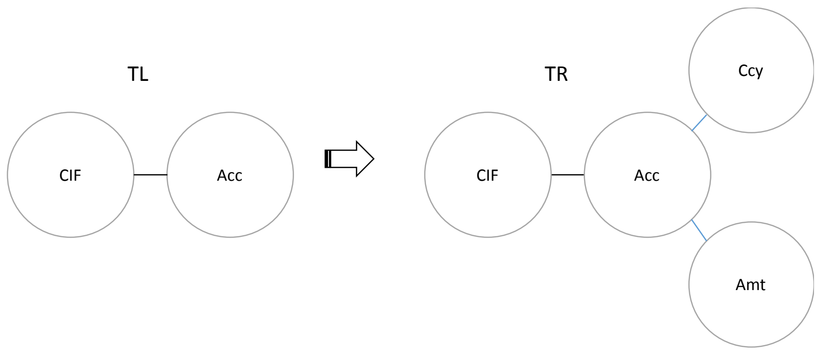


Fig. 7. Graph production on the target model

In the above example, we can consider productions SL→SR and TL→TR as compatible, provided that we can map the names of the attributes between the models. Such a mapping needs to be created by an expert during the first transformation of given model and recorded as meta-data that can be used for future automated transformations.

The simple example above presents transformation from S → T. At this point, we can assume that there is another model (e.g. X, Y, Z), compatible with S and T, but having different structures and using different names of the attributes. These models can be based e.g. on country regulations and local languages. In order to make the model universal, we would need to look for compatible productions, using a pre-defined set of meta data. Such a mapping would require each pair of the models to be compared and individually transformed, i.e.:

S → X, S → Y, S → Z, X → T, Y → T, Z → T.

In order to simplify this process, we can introduce an intermediary model that can be used as a source and target for each model being analyzed for suitability for transformation. The intermediary model would be called a Pattern Model (P). Provided that there are a large number of models, this approach would significantly limit number of potential model → model transformations. Resulting in a list of transformations that would look like the following:

S ←→ P, X ←→ P, Y ←→ P, Z ←→ P.

The following example demonstrates two productions, resulting in different target models:

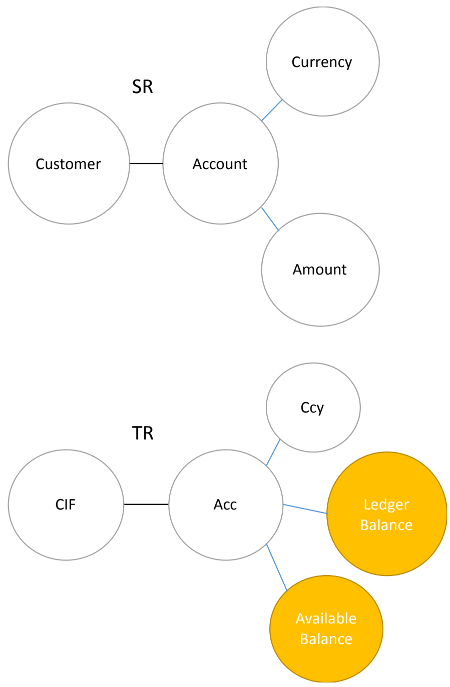


Fig. 8. Example of productions, resulting in different target models

8. PRACTICAL APPLICATIONS OF THE DATA MODEL TRANSFORMATIONS

Automation of the Data Model Transformation Processes becomes a critical success factor in multiple commercial solutions. Listed below are practical examples where a “higher than current” level of automation would significantly improve both timing and cost efficiency.

8.1. Integration Projects

The EU’s Payment Services Directive (PSD) was originally published in December 2007. The recently proposed PSD2 introduces Trusted Third Party Account Access, which are represented by two acronyms: Third Party Payment (TPP) under the Access to Accounts (XS2A) rule. Besides legal and operational discussions, XS2A has significant impact on how Banks will enable TPP providers to access their resources (Skinner, 2015).

Proprietary data formats would require each payment provider to commit to the API exposed by particular banks. On the other hand, banks my want particular payment providers to easily access their resources; in this case, banks may need to adopt the API exposed by the TPP.

As discussed in the previous chapters, particular data formats might significantly vary, while the amount of information is very similar between particular parties.

8.2. Advanced Analytics

The advent of Big Data, supported by virtually unlimited storage capacities available at the low cost, results in very large amounts of data stored by particular organizations. Unlike typical database solutions, Big Data doesn’t require data formats to be precisely defined and followed. It is important to collect a lot of data and use proper analytical tools to make meaningful use of that data.

8.3. Automation of Test Data Preparation

Automated testing is one of the key components of currently emerging DevOps. The term “DevOps” is a compound of “development” and “operations”, and is a movement or practice that emphasizes the collaboration and communication of both software developers and other information-technology (IT) professionals, while automating the process of software delivery and infrastructure changes (Cortet, 2014). An important part of DevOps is the automation of testing processes. While in traditional testing approaches, test data are prepared together with test scripts; an agile approach, combined with DevOps, requires quick access to the test data for new developments, where actual data doesn’t yet exist. On the other hand, test data can be created based on some other data, that are similar from a content perspective but are completely different from format perspective. Once again, the automated transformation of such a data would significantly improve performance of the projects.

9. CONCLUSIONS

Transformation of a data model to its formal graph representation is necessary, in order to allow further automation of the transformation process. Provided that each of the source data models go through this transformation, the resulting models will share common set of characteristics, such as attribute names and basic data structures. Resulting high level of automation of the transformation process would provide significant savings, both in terms of time and money, compared to similar work performed by a human expert.

REFERENCES

ADRM Software. (2018, August 1). *Business Area Data Models*. Retrieved from http://www.adrm.com/data-model-business-area.html

Angles, R., & Gutierrez, C. (2008). *Survey of graph database models*. New York, USA: ACM.

Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., & Yergeau, F. (2008). *Extensible Markup Language (XML) 1.0* (Fifth Edition). Retrieved from https://www.w3.org/TR/xml

Cortet, M. (2014). *Access to the Account (XS2A): accelerating the API-economy for banks?* Retrieved from https://innopay.com/blog/access-to-the-account-xs2a-accelerating-the-api-economy-for-banks

Holman, K. (2018). *OASIS Universal Business Language (UBL) TC*. Retrieved from https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ubl

ISO 20022. (2018, August 1). *Universal financial industry message scheme*. Retrieved from https://www.iso20022.org

Kotulski, L. (2013). *Rozproszone transformacje grafowe*. Kraków, Poland: Wydawnictwo AGH.

McKnight, W. (2014). *IBM Industry Data Models in the Enterprise*. Retrieved from https://www-01.ibm.com/software/data/industry-models/

Roman, D. (2006). *Canonical Data & Process Models for B2B Integration*. Retrieved from http://ceur-ws.org/Vol-170/paper3.pdf

Skinner, Ch. (2015). *How will Banks organise themselves to manage APIs built for PSD2/XS2A?* Retrieved from http://thefinanser.com/2015/11/how-will-banks-organise-themselves-to-manage-apis-built-for-psd2-xs2a.html/

Sleger, G. (2010). *Data Transformation Mapping – Can it be Automated?* Retrieved from https://www.cleo.com/blog/data-transformation-mapping-can-it-be-automated

SWIFT. (2018, August 1). *Financial messaging services*. Retrieved from https://www.swift.com/about-us/discover-swift/messaging-standards

Thompson, H. & Lilley, C. (2014). *XML Media Types, RFC 7303*. Retrieved from https://tools.ietf.org/html/rfc7303