

Izabela KUTSCHENREITER-PRASZKIEWICZ^{1*},
Jerzy JEDRZEJEWSKI²

A METHODOLOGICAL APPROACH TO ASSEMBLY TIME STANDARD ESTIMATION BASED ON INCOMPLETE CHARACTERISTICS OF THE PRODUCTION PROCESS AND USING SMALL DATABASES

The problem solved in this article concerns assembly planning, which is time-consuming, but crucial in the development of mechanical products. At the product design stage there is no complete information about the manufacturing process, so it is necessary to develop an approach to help process the uncertain and incomplete information. In order to compare different product variants, the assembly time standard has to be estimated on the basis of the incomplete product and production process characteristics. This paper presents a method for estimating the assembly time standard using time classes, decision tree and evidence theory.

1. INTRODUCTION

As mechanical assemblies are complex structures [1], their planning is challenging. Using the assembly design concept [2] one can select the best construction solution for the designed product, taking into account, i.a., the estimated standard assembly time. At an early stage of product development, data on the manufacturing process are not fully known. Moreover, there is a gap in the methodology for determining the duration of production operations.

The existing time determination methods require a lot of data on the product and the production process. Therefore it is vital to develop a time determination method which will take into account the missing data. A theory which proves useful for processing incomplete information is the evidence theory, known as the Dempster-Shafer theory (DST). The approach proposed in this paper combines machine learning methods, such as decision tree induction, with a time standard estimation method using time classes. The proposed approach employs training set data and DST-based uncertainty reduction and focuses on estimating assembly time, which is done early in product development. The existing methods for

¹ Faculty of Mechanical Engineering and Computer Science, University of Bielsko-Biala, Poland

² Department of Machine Tools and Mechanical Technologies, Wrocław University of Science and Technology, Poland

* E-mail: ipraszkiewicz@ubb.edu.pl

<https://doi.org/10.36897/jme/190313>

determining time standards are inadequate for predicting time standards at an early stage of product development. Methods for determining time standards are classified into work measurement techniques, historical records and estimations (e.g. based on time classes TC) [3]. Work measurement techniques include: direct time study, predetermined motion time systems (e.g. MTM), standard data systems (e.g. time normative TN) and work sampling. Using these methods one can measure or calculate work, but the methods are inappropriate for planning assembly tasks with uncertain and incomplete data. Chen et al. [4] mention that the existing work time estimation methods are based on expert experience and historical data. Another approach was presented by Kim et al. [5] who used motion studies and simulations to achieve optimal job performance. Bentaha et al. [6] discuss the disassembly process and use the Monte Carlo simulation to deal with uncertainty in disassembly line design, treating task time as a random variable with a known probability distribution. Somala et al. [7] apply the machine learning approach to the time period estimation of masonry infilled reinforced concrete frames. Zhang et al. [8] use a neural network to forecast the human operator's motion during the assembly process. Kwon et al. [9] apply the genetic algorithm (GA), the multiple linear regression analysis (MLR), the feature counting method and the fuzzy-analytical hierarchy process to case-based reasoning for predicting repair time. Assembly modelling and time estimation during an early phase of assembly system design is discussed by Salmi et al. [10] who analyse the appropriate level of automation for the design and optimisation of new assembly systems. As in the early stages of product development some data are missing, it is necessary to use methods which can cope with uncertain data. Since the accuracy of the time standard estimation depends on the estimator's evaluation [3], it is necessary to use, e.g., machine learning methods to obtain more precise person-independent data. In order to fill the gap in the time standard setting methods, it is vital to develop a method which will be independent of human experience and will be able to predict task time even when all the characteristics of the assembly process are unknown.

The evidence theory provides a framework helpful in representing and processing uncertain information [11]. Razavi et al. [12] used the k-Nearest Neighbours algorithm, the Expectation–Maximization algorithm and the DST theory as a method for estimating the final results of missing data imputation in machine learning. Ma et al. [13] present approaches to a partial classification based on the DST of belief functions. Guo et al. [14] use DST to support risk assessment, fusing multi-source information and taking into consideration uncertainty, conflicts and dynamics. Deng et al. [15] introduce probabilistic information content concepts useful for decision making. Du et al. [16] analysed a group decision-making process and proposed a group inference method based on DST. Yu et al [17] used DST to avoid clustering errors, whereas Strat [18] constructs a decision tree using DST. There are different approaches to assembly planning. Krist et al. [19] presented a method in which individual experts' knowledge is integrated into planning the assembly process. Pimminger et al. [20] proposed a General Assembly Task Model (GATM) for structuring an assembly task. Ghadge et al. [21] presented an approach to knowledge acquisition and management in the manual assembly process. DST and the decision tree have been applied to decision-making related to the selection of production tools [22].

There is a gap in the methods for estimating the time standard at an early stage of product development. The methods presented in the literature focus on the analysis of assembly tasks

on the shop floor. An innovative combination of the machine learning method in the area of training set preparation with DST offers new opportunities. This paper proposes a way to avoid unknown data in the process of estimating the assembly time standard at an early stage of product development. The assembly time standard depends on various attributes, including, i.a., the characteristics of the parts to be assembled, the assembly method, the tools used, the workstation layout and the feeding of parts. Since some of these data are not known at an early stage in the development of a product part, it is necessary to adopt an approach which helps to avoid unknown data on assembly characteristics.

2. THE PROPOSED APPROACH

The proposed approach (Fig. 1) focuses on predicting the assembly time standard (ATS) based on classes of time standards, graph theory and DST and requires the following assumptions: the assembly process is divided into separate assembly tasks, the worker is authorized to perform the task, has sufficient experience and skills, and works at a normal pace, the workstation, equipment and tools meet quality requirements, the workstation meets health and safety requirements, the input material, information and energy are of standard quality, the output item meets quality requirements.

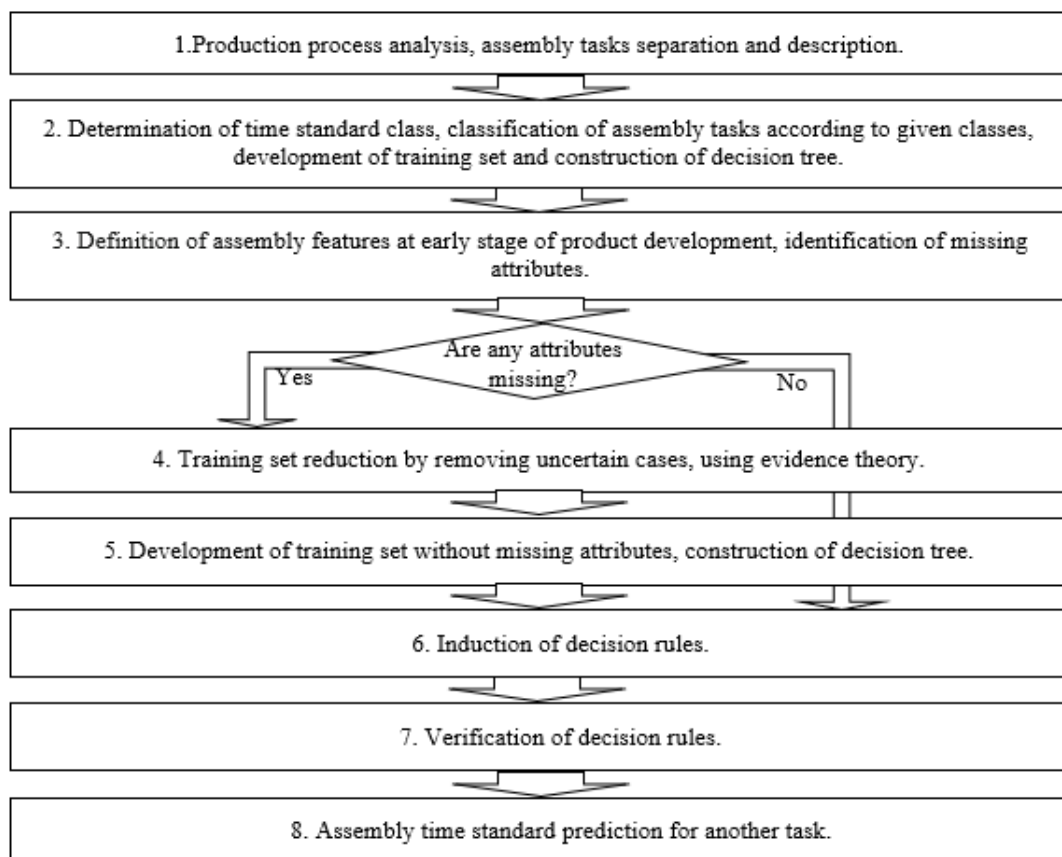


Fig 1. The proposed approach to estimating standard assembly time

The first stage of the proposed approach presented in Fig. 1 concerns the analysis of the production process and the separation and description of the assembly tasks. The assembly task can be classified according to the idea of group technology. The assembly task can be analysed according to the concepts of the work system, which include: task to be performed – the purpose of the work system, input – material and information, worker – experience, qualifications, worker's well-being, workstation and tools – type, parameters and technical condition of the workstation, result – the produced object, environment – temperature, pollution, humanity, human relations, etc., work method – sequence of tasks, work elements and basic elements of movement. The assembly task can be divided into a work element and a basic motion element. The assembly process should be characterised for each component to be assembled, specifying the assembly sequence and the required tools. The assembly process can be divided into work tasks and analysed by characteristics, including parameters such as dimensions, mass, distances, etc. In the proposed method, structural and technological features affecting assembly time are analysed. Dimensional accuracy is treated as a feature which determines assimilability and is not considered in the proposed method.

The second stage of the proposed approach includes the determination of the time standard class, the classification of the assembly tasks according to the given classes, the development of a training set and the construction of a decision tree. The time standard determination method used in the proposed approach is based on time standard classes, including work tasks and fixed values of labour-consumption standards. A class catalogue of typical assembly tasks can be formed on the basis of the equation (1) [23].

$$w_j = \frac{e \sqrt{12T_m(t_n)_j}}{196} \quad (1)$$

where w_j represents the width of *class j*, e - accuracy at the 95% confidence level, T_m - the reference period, $(t_n)_j$ - the time standard for *class j*. Each typical assembly task should be assigned to an appropriate class characterised by a lower bound, an upper bound and an average value used as a fixed time standard value for the given class.

The proposed approach uses a supervised learning method which requires a training and validation sets (see [24]) containing examples of work task characteristics and an exemplary time standard class as the basis for decision making in the decision process. Decision tree induction is used as a tool in machine learning. The attributes characterizing the components to be assembled, the workstations and the tools can be used in the training process. Some of the attributes can be used in a decision tree and the other can be omitted. The structure of the decision tree depends on the applied algorithm (e.g. the ID3 algorithm).

During the third stage of the proposed approach the assembly features at an early stage of product development are defined and the missing attributes are identified. If the decision tree contains attributes which are not known in the product design phase, the training set must be reconstructed. The number of attributes which are unknown at an early stage of product development should be reduced. It may happen that after the training set is reduced, the training example is not unambiguously classified.

The fourth stage of the proposed approach involves reducing the training and validation sets by removing uncertain cases using DST. Some of the attributes in the assembly tasks are unknown at an early stage of product development. Thus the data used to estimate the time

standard are incomplete and uncertain. A method which seems promising for handling uncertain data is DST. This method focuses on managing uncertain information from multiple sources [16]. According to DST, the basic probability assignment (BPA) function of two pieces of evidence m_1 and m_2 can be calculated from the equation (2) [16].

$$m_{12}(\theta) = [m_1 \oplus m_2](\theta) = \frac{\sum_{B \cap C = \theta} m_1(B)m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B)m_2(C)} \quad (2)$$

where $m_1 \oplus m_2(\theta)$ is the fused belief degree of the two pieces of evidence; $1 - \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ is the degree of contradictory information between B and C , referred to as a normalization factor; $\theta = \{\theta_n | n = 1, \dots, N\}$ is the frame of discernment.

DST uses a belief function, which represents the belief involved solely in the hypothesis, and a plausibility function which represents the belief potentially supporting the hypothesis [25]. The belief function (Bel) and the plausibility function (Pl) of subset θ_n of Θ can be defined on the basis of the equation (3) [16].

$$\begin{cases} Bel(\theta_n) = \sum_{B \subseteq \theta_n} m(B) \\ Pl(\theta_n) = \sum_{\theta_n \cap B \neq \emptyset} m(B) \end{cases} \quad (3)$$

where $[Bel(\theta_n), Pl(\theta_n)]$ is a confidence interval indicating the degree of certainty of θ_n . In order to calculate Bel and Pl one needs data coming from the production process. In the proposed approach, pieces of evidence are understood as probability which comes from different periods. In the given periods the number of cases from the analysed variants is included in the probability calculations and then Bel and Pl are determined. The duration of periods depends on the production process and the manufacturing orders.

The fifth stage of the proposed approach involves developing a training and validation sets without missing attributes and building a decision tree. Uncertain cases with the same input characteristics and different output classifications are compared and cases with a lower Bel value are removed from the training set. The ambiguously classified variants are compared for the selected periods. If two cases have the same Bel value, the case with the higher Pl value is removed from the training set. If both Bel and Pl are the same, then the case with the lower time class should be removed from the training set. A decision tree can be built using one of the well-known algorithms, such as ID3.

The sixth stage focuses on the induction of decision rules, which can be taken directly from the decision tree (each path in the decision tree represents a decision rule). Attributes in the graph nodes and values in the branches represent the rules' premises while the decision leaves represent conclusions in the decision rules.

The seventh stage validates the decision rules with a testing set of examples according to cross-validation approach. Verification and validation of the manufacturing process has been discussed by Maropoulos and Ceglarek [26] and Ebrahimi Araghizad et. al. [27].

The eighth stage uses the decision rules to predict the assembly time standard for another task.

The proposed approach can be used to determine the assembly time of a typical subassembly on the basis of typical product and process parameters and a typical workstation and tool.

3. EXPERIMENTAL STUDY

The aim of this experimental study was to create decision rules for predicting the time standard of bearing assembly, which is one of the key components of a gearbox that affects its reliability.

The assembly process of a bearing shaft in a gearbox subassembly, for which *ATS* had been calculated using the time class method (Table 1), was analysed. The time class method involves determining time by measurement methods such as time study or elementary motion analysis and then assigning the time to a time class. The average value of the time class is the standard value of the time of the analysed task. In order to apply the time class method, it is necessary to determine class intervals. The presented example involves the following assumptions: the assembly of ball bearings with a bearing bore diameter of 50-210 mm has been analysed (the determined time values are shown in the fourth column of Table 2), T_m - a reference period equal to 40 h, e - accuracy equal to 2%. The $(t_n)_j$ value in equation (1) for class one is calculated as the minimum time in the set of the analysed cases minus 30%. For the subsequent classes the upper bound of the class becomes the lower bound of the next time class. Thus, each task can be unambiguously classified into a particular time class.

Table 1. Definition of *ATS* classes

Class	(Lower, Upper> bound of class [h])	Average value of class [h]
A	(0.08, 0.14>	0.11
B	(0.14, 0.22>	0.18
C	(0.22, 0.33>	0.28
D	(0.33, 0.46>	0.39

Table 2. Training set with full characteristics

Attribute (<i>Bid</i>)	Attribute (<i>Hc</i>)	Attribute (<i>H</i>)	<i>ATS</i> [h]	Class
60	0	No	0.11	A
80	6	Yes	0.15	B
50	2	Yes	0.14	A
50	6	Yes	0.16	B
200	6	Yes	0.20	B
210	2	Yes	0.19	B
70	6	Yes	0.17	B
205	6	Yes	0.20	B
212	2	Yes	0.19	B
55	2	Yes	0.14	A

In order to create a decision tree, it is necessary to build a training set containing bearing assembly time standards using time classes as decision variables with assembly product and process characteristics as the input characteristics.

The assembly characteristics in the training set (Table 2) apply to the bearing assembly and include attributes such as: the bearing internal diameter (Bid) depending on the bearing size; simultaneous bearing heating (Hc) depending on the organization of the assembly process; heating (H) – a technological factor recommended for the assembly process of some bearings. On the basis of the data contained in Table 2 a decision tree was created (Fig. 2).

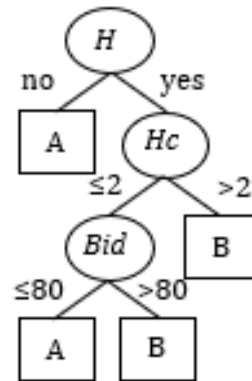


Fig 2. Decision tree for ATS determination with full characteristics

In an enterprise it may happen that the analysed information does not cover the entire range of data. Therefore a training set with reduced uncertainty needs to be built as described above.

DST is helpful in combining information from previous analytical periods, relating to uncertain cases in the training set. A universal training set can be created for a typical subassembly based on a typical manufacturing process, but uncertainty reduction must take into account the company's experience. Therefore the training set is created taking into account this experience and an analysis is carried out in accordance with the proposed method. Since at an early stage of product development some organizational and technological solutions are unknown, Table 3 contains an uncertain training set with a limited number of attributes. The internal diameter of the bearing is the only attribute known at an early stage of product development and the reduction in the number of the other attributes causes uncertainty in the training set. Uncertain cases in the training set are cases 3 and 4. DST was used to reduce the uncertainty.

Table 3. Training set with uncertain cases

Case no	Attribute (Bid)	Class
1	60	A
2	80	B
3	50	A
4	50	B
5	200	B
6	210	B
7	70	B
8	205	B
9	212	B
10	55	A

Table 4 shows the probability connected with the application of the two variants: A and B in the four periods: from $m1$ to $m4$. The confidence intervals determining the degree of certainty of the variants are presented in Table 5. The training set with reduced uncertainty is presented in Table 6. The corresponding decision tree is shown in Fig. 3.

Table 4. Probability assessment

Periods \ Variants	$m1$	$m2$	$m3$	$m4$
A	0.3	0.1	0.4	0.3
B	0.4	0.3		0.2

Table 5. Fused belief degree

Variants	Bel	Pl
$m1234(A)$	0.52	0.61
$m1234(B)$	0.39	0.48

Table 6. Training set with reduced uncertainty

Case no	Attribute (Bid)	Class
1	60	A
2	80	B
3	50	A
4	50	A
5	200	B
6	210	B
7	70	B
8	205	B
9	212	B
10	55	A

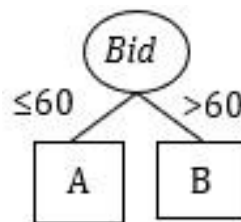


Fig 3. Decision tree with reduced uncertainty for ATS determination

Using the proposed approach all the cases in the training set (Table 6) are classified without error. On the basis of the decision tree the following decision rules are created:

$$\text{IF } Bid \leq 60, \text{ then } A \quad (4)$$

$$\text{IF } Bid > 60, \text{ then } B \quad (5)$$

By applying the proposed approach to the test set comprising examples of the assembly of bearings with an internal diameter of 70 and 40 mm the unit assembly time was determined. It amounted to respectively 0.18 and 0.11 hours (average time values for class B and A). In addition the cross validation gives good results.

The proposed approach yields positive results. It reduces classification errors and can be used in early product development for a given assembly process in the expert system.

The proposed approach takes into account unitary manufacturing process characteristics relating to the manufacture of products differing from one another.

A comparison of the time standard setting methods such as: MTM (1), time normative TN (2), time classes TC (3) and proposed approach (4) is shown in Table. 7.

Table 7. Comparison of time standard data prediction

Sample No	Attribute (Bid)	ATS data prediction [h]			
		Method (1)	Method (2)	Method (3)	Method (4)
1	60	0,11	missing	0,11	0,11
2	80	0,15	0,17	0,18	0,18
3	50	0,14	0,16	0,11	0,11
4	50	0,16	0,17	0,18	0,11
5	200	0,2	0,2	0,18	0,18
6	210	0,19	0,22	0,18	0,18
7	70	0,17	0,17	0,18	0,18
8	205	0,2	0,23	0,18	0,18
9	212	0,19	0,22	0,18	0,18
10	60	0,14	0,16	0,11	0,11

6. CONCLUSION

Assembly time standard estimation can be supported by a machine learning method such as the decision tree and DST. At an early stage of product development it is possible to use assembly features relating to the assembled components, but some attributes describing the assembly process must be omitted. Thus, the training set used to create the decision trees contains uncertain and incomplete data.

DST can be employed as a method of processing this type of data. The presented approach uses decision tree induction based on the training set and takes into account the changes in the historical data in the given periods. Ultimately, a decision tree is built and decision rules are induced from it.

The proposed approach can be applied to an enterprise which has experience in assembling a particular class of subassemblies and it is helpful in predicting the assembly time at the product design stage, as well as in planning the production process.

An example of the application of the proposed approach to determining assembly time is presented based on the assembly process of gearbox bearings, which is one of the key processes affecting the reliability of gearboxes.

The approach presented is generic and can be applied to any type of assembly, and the features in the training set depend on the component being assembled and the tools used. The case study presented is a numerical example focusing on explaining the proposed method.

Future research will focus on finding features that are important for creating a training set for different types of assembly.

REFERENCES

- [1] BONINO B., RAFFAELI R., MONTI M., GIANNINI F., 2021, *A Heuristic Approach to Detect CAD Assembly Clusters*, *Procedia CIRP*, 100, 463–468, doi.org/10.1016/j.procir.2021.05.105.
- [2] EZPELETA I., PUJANA U., ISASA I., AYERBE J., JUSTEL D., 2021, *New Design for Assembly (Dfa) Methodology for Large and Heavy Parts Assembled on Site*, *Procedia CIRP*, 100, 145–150, doi.org/10.1016/j.procir.2021.05.078.
- [3] GROOVER M., 2016, *Work Systems, the Methods, Measurements and Management of Work*, Pearson.
- [4] CHEN J., ZHOU D., KANG L., MA L., GE H., 2020, *A Maintenance Time Estimation Method Based on Virtual Simulation and Improved Modular Arrangement of Predetermined Time Standards*, *International Journal of Industrial Ergonomics*, 80, 103042, doi.org/10.1016/j.ergon.2020.103042.
- [5] KIM J., GOLABCHI A., HAN S., LEE D., 2021, *Manual Operation Simulation Using Motion-Time Analysis Toward Labour Productivity Estimation: a Case Study of Concrete Pouring Operations*, *Automation in Construction*, 126, 103669, doi.org/10.1016/j.autcon.2021.103669.
- [6] BENTAHA M.L., BATAIA O., DOLGUI A., HU S.J., 2014, *Dealing With Uncertainty in Disassembly Line Design*, *CIRP Annals – Manufacturing Technology*, 63, 21–24, doi.org/10.1016/j.cirp.2014.03.004.
- [7] SOMALA S., KARTHIKEYAN K., MANGALATHU S., 2021, *Time Period Estimation of Masonry Infilled RC Frames Using Machine Learning Techniques*, *Structures*, 34, 1560–1566, doi.org/10.1016/j.istruc.2021.08.088.
- [8] ZHANG J., LIU H., CHANG Q., WANG L., GAO R.X., 2020, *Recurrent Neural Network for Motion Trajectory Prediction in Human Robot, Collaborative Assembly*, *CIRP Annals – Manufacturing Technology*, 69, 9–12, doi.org/10.1016/j.cirp.2020.04.077.
- [9] KWON N., AHN Y., SON B., MOON H., 2021, *Developing a Machine Learning-Based Building Repair Time Estimation Model Considering Weight Assigning Methods*, *Journal of Building Engineering*, 43, 102627, doi.org/10.1016/j.jobe.2021.102627.
- [10] SALMI A., DAVID P., BLANCO E., SUMMERS J., 2015, *Assembly Modelling and Time Estimating During the Early Phase of Assembly Systems Design*, *IFAC-Papers On Line*, 48/3, 81–87, doi.org/10.1016/j.ifacol.2015.06.062.
- [11] GENG X., LIANG Y., JIAO L., 2021, *ARC-SL: Association Rule-Based Classification with Soft Labels*, *Knowledge-Based Systems*, 225, 107116, doi.org/10.1016/j.knosys.2021.107116.
- [12] RAZAVI-FAR R., CHENG B., SAIF M., AHMADI M., 2020, *Similarity-Learning Information-Fusion Schemes for Missing Data Imputation*, *Knowledge-Based Systems*, 187, 104805, doi.org/10.1016/j.knosys.2019.06.013.
- [13] MA L., DENOEU T., 2021, *Partial Classification in the Belief Function Framework*, *Knowledge-Based Systems*, 214, 106742, doi.org/10.1016/j.knosys.2021.106742.
- [14] GUO K., ZHANG L., 2021, *Multi-Source Information Fusion for Safety Risk Assessment in Underground Tunnels*, *Knowledge-Based Systems*, 227:107210, doi.org/10.1016/j.knosys.2021.107210.
- [15] DENG Z., WANG J., 2020, *A Novel Decision Probability Transformation Method Based on Belief Interval*, *Knowledge-Based Systems*, 208, 106427, doi.org/10.1016/j.knosys.2020.106427.
- [16] DU Y., ZHONG J., 2020, *Group Inference Method of Attribution Theory Based on Dempster-Shafer Theory of Evidence*, *Knowledge-Based Systems*, 188, 104985, doi.org/10.1016/j.knosys.2019.104985.
- [17] YU H., CHEN L., YAO J., 2021, *A Three-Way Density Peak Clustering Method Based on Evidence Theory*, *Knowledge-Based Systems*, 211, 106532, doi.org/10.1016/j.knosys.2020.106532.
- [18] STRAT T., 1990, *Decision Analysis Using Belief Functions*, *International Journal of Approximate Reasoning*, 4, 391–417.
- [19] KRIST K., SIEVERS T., ONKEN A., KODJO Y., TRACHT K., 2020, *Application of Derivative Products for Integrating Expert Knowledge Into Assembly Process Planning*, *Procedia CIRP*, 88, 88–93, doi.org/10.1016/j.procir.2020.05.016.
- [20] PIMMINGER S., KURSCHL W., PANHOLZER L., NEUMAYR T., AUGSTEIN M., ALTMANN J., HEINZELREITER J., 2020, *Assembly Task Analysis Using the General Assembly Task Model (GATM) on the Shop Floor*, *Procedia CIRP*, 93, 1109–1114, doi.org/10.1016/j.procir.2020.04.007.
- [21] GHADGE K., CHAKRABARTI A., 2020, *A Framework for Knowledge Management in Manual Assembly Processes*, *Procedia CIRP*, 88, 94–97, doi.org/10.1016/j.procir.2020.05.017.
- [22] KUTSCHENREITER-PRASZKIEWICZ I., MATUSZNY M., 2022, *Knowledge Base Development for Assembly Planning Using Evidence Theory*, *Journal of Machine Engineering*, 22, 1–17, doi.org/10.36897/jme/149185.

- [23] WOLK R., STRZELECKI T., 1993, *Study of Methods and Standardization of Work*, (in Polish), Warsaw University of Technology Publishing House, Warsaw.
- [24] WITTEN I., FRANK E., HALL M., 2011, *Data Mining Practical Machine Learning Tools and Techniques*, Morgan Kaufmann.
- [25] BEECHEY M., KYRIAKOPOULOS K., LAMBOTHARAN S., 2021, *Evidential Classification and Feature Selection for Cyber-Threat Hunting*, Knowledge-Based Systems 226 107120, doi.org/10.1016/j.knosys.2021.107120.
- [26] MAROPOULOS P.G., CEGLAREK D., 2010, *Design Verification and Validation in Product Lifecycle*, CIRP Annals – Manufacturing Technology, doi.org/10.1016/j.cirp.2010.05.005.
- [27] EBRAHIMI ARAGHIZAD A, TEHRANIZADEH F., KILIC K., BUDAK E., 2023, *Smart Tool-Related Faults Monitoring System Using Process Simulation-Based Machine Learning Algorithms*, Journal of Machine Engineering, 23, 4, 18–32, <https://doi.org/10.36897/jme/174018>.