

## Product Manufacturing Process Similarity Measure Based on Attributed Graph Matching

Yongjian Zhang<sup>1,a\*</sup>, Jinshan Liu<sup>2,b</sup> and Lin Wang<sup>1,c</sup>

<sup>1</sup> School of Naval Architecture and Ocean Engineering Harbin Institute of Technology (Weihai),  
Weihai, Shandong, 264209, China

<sup>2</sup> Beijing Satellite Manufacturing Factory, Beijing, 100094, China

<sup>a</sup> zhangyj@hitwh.edu.cn, <sup>b</sup> secularbird@263.net, <sup>c</sup> wanglin105@126.com

**Keywords:** manufacturing process; similarity measure; attributed graph

**Abstract.** To support implicit knowledge acquisition in manufacturing process design and increase the enterprise's rapid response capabilities and competitiveness, similarity measure on manufacturing process was studied in this paper. Based on the characteristic description of modular process unit, process representation model using attributed graph was proposed. Manufacturing process similarity measure method considering both unit similarity and sequence similarity on optimal matching of weighted bipartite graph was given. Application of the method using on building process platform of satellite product is introduced and discussed.

### Introduction

Nowadays time for design and production need to be decreased to increase the enterprise's rapid response capabilities and competitiveness. Mining implicit knowledge from history design cases and instances become more and more important. In manufacturing design phase, concepts of process platform and process family were proposed and methods based on product family were studied<sup>[1-3]</sup>. To building the process family and platform, process representation model and measurement that quantifies the similarity between manufacturing processes based on the representation need to be studied firstly<sup>[4]</sup>. Many similarity measure methods have proposed, most of them focus on common datatype like tree, graph, net, and only considering structure of data model<sup>[5,6]</sup>, measurement on manufacturing process need further research. This paper is organized as follows: In the next section, manufacturing process representation using attributed graph is given in detail. Next, on the basis of process graph representation, similarity measure method based on graph matching is introduced. In application section, calculation example is given. In the last section, the paper is concluded.

### Manufacturing Process Representation

**Manufacturing Process.** Manufacturing process is a combination of a series of orderly processing units. Process units like procedure, step and activity are modular components of the process. Process units generally carried out in form of work type, processing content varied with different type, and units with different type has no comparability. Even with same work type, due to the different manufacturing requirements, differences of owned resources, and process designer's experience personality and habits, process units still make a big difference. If two processes contains same kind and same number of units, also with the same order of units, the two processes are the same.

**Attributed Graph Based Process Representation.** Formally, an attributed graph denoted as  $G = (U, S, UA, F, A)$  is used to describe all information of manufacturing process mentioned before. For graph  $G$ ,  $U$  is the set of graph vertices which represent process units;  $S \subseteq U \times U$  is the set of graph edges joining two distinct process units which indicates sequence relationship between units;  $UA$  is the set of manufacturing process unit attributes;  $F$  is a function  $U \rightarrow UA$  which assigns attributes to related process unit;  $A$  is a set of attributes that can be associated with any process unit in  $U$ .

Manufacturing process has two kinds of unit sequence: serial sequence and parallel sequence. Serial sequence illustrates successively relationship on time between coupled units. Parallel sequence

represents fork and merge logic in the process. As is shown in Fig.1, process unit  $U_i$  followed by  $U_j$  form a serial sequence. Process unit  $U_m$  and  $U_n$  between  $U_k$  and  $U_l$  are paralleled unit, sequence  $(U_k, U_m)$  and  $(U_k, U_n)$  represents fork logic, while sequence  $(U_m, U_l)$  and  $(U_n, U_l)$  represents merge logic.

For each process unit  $U_i \in U$ , attributes  $UA$  can be described in detail as  $UA = (type, level, MA, FA, R)$ . In  $UA$ ,  $type$  is work type of process unit with value of turning, milling, assembly, etc. which distinguish one unit from others roughly. The value of  $type$  is determined by company standard;  $level$  describes the importance degree of process unit with value from 0 to 1. In a specific process, units having bigger  $level$  value are core unit which highlight process characteristics, while units having small  $level$  value are auxiliary unit which refine process in detail and keep process's integrity;  $MA = [a_1 \dots a_{MJ}]$  is the set of managed attributes such as name, description, create time, etc.;  $FA = [p_1 \dots p_{FJ}]$  is the set of featured attributes associated with  $type$  of process unit which highlight process unit's characteristics. For example, process unit with turning type has featured attributes such as turning speed, cutting dosage, cutting force, etc.  $R = [r_1 \dots r_{RJ}]$  is the set of resources used by process unit.

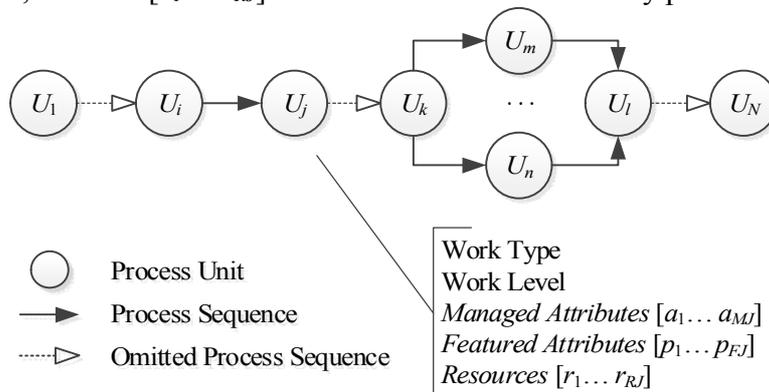


Fig. 1 Attributed Graph Based Process

### Similarity Measure Based on Graph Matching

Similarity between manufacturing processes is measured from two aspects: process unit similarity and process sequence similarity. Because sequence indicating relationship between units, process sequence similarity should be measured on the basis of process unit similarity. In this paper, a similarity measure method using process units optimal matching is given. Giving two manufacturing processes  $P_a$  and  $P_b$ , the similarity denoted as  $sim(P_a, P_b)$  is achieved using following methods.

**Process Unit Similarity Measure.** Featured attributes show process unit's characteristics, so similarity of the  $i$ th unit  $U_{ai}$  from process  $P_a$  and the  $j$ th unit  $U_{bj}$  from process  $P_b$  denoted as  $sim_U(U_{ai}, U_{bj})$  is achieved by evaluating similarity of  $FA_{ai}$  and  $FA_{bj}$ . Since featured attributes is associated with  $type$  of process unit, only units with same type value are comparable. Most of engineering parameters is of quantitative nature. The value of the parameters can be expressed as a number. A similarity measure is a function  $sim \rightarrow [0, 1]$ , Eq. 1 is used for transforming a distance measure into a similarity measure, in which,  $K$  is the quantity of featured attributes, Cosine distance or Euclidean distance can be used for calculating  $dist(p_{aik}, p_{bjk})$ .

$$sim_U(U_{ai}, U_{bj}) = \frac{1}{K} \sum_{k=1}^K \frac{1}{(1 + dist(p_{aik}, p_{bjk}))} \quad (1)$$

**Process Units Matching.** Process  $P_a$  and  $P_b$  may have different quantity and different types of units, to measure the similarity of  $P_a$  and  $P_b$ , units form different process should be compared, and units matching should be firstly achieved.

Process units matching is to find the matching set  $M$  and make  $M$  meet the conditions as shown in Eq. 2. This problem can be converted into an optimal matching problem in the weighted complete bipartite graph  $BG = (X, E, Y)$ , where  $X, Y$  are the set of process units from  $P_a$  and  $P_b$  with filtering

condition of unit *level* value not less than the given threshold  $\epsilon$ ,  $E$  is the set of links between units from  $X$  and  $Y$ , and  $|X| = |Y| = N$ . To meet  $|X| = |Y|$ , virtual units is added to units set  $X$  or  $Y$ . A link  $e = (U_{ai}, U_{bj}) \in E$  with the value of  $sim_U(U_{ai}, U_{bj})$  which is defined in Eq. 1, and  $sim_U(U_{ai}, U_{bj})=0$  when  $U_{ai}$  or  $U_{bj}$  is virtual unit added before. Optimal matching problem can be solved using Kuhn-Munkras (KM) algorithm.

$$\sum_{(U_{ai}, U_{bj}) \in M} sim_U(U_{ai}, U_{bj}) \geq \sum_{(U_{ai}, U_{bj}) \in M'} sim_U(U_{ai}, U_{bj}), \quad \forall M' \quad (2)$$

**Process Similarity Measure.** On the basis of process units matching, similarity of process  $P_a$  and  $P_b$  denoted as  $sim(P_a, P_b)$  can be measured by combining node (unit) similarity  $sim_N(P_a, P_b)$  and sequence similarity  $sim_S(P_a, P_b)$  using weighted summing as is shown in Eq. 3. In this equation,  $sim_N(P_a, P_b)$  is the integration of unit similarity from the optimal units matching given in Eq. 2, which is shown in Eq. 4.

$$sim(P_a, P_b) = w sim_N(P_a, P_b) + (1 - w) sim_S(P_a, P_b) \quad (3)$$

$$sim_N(P_a, P_b) = \frac{2}{|P_a| + |P_b|} \sum_{(U_{ai}, U_{bj}) \in M} sim_U(U_{ai}, U_{bj}) \quad (4)$$

Sequence similarity  $sim_S(P_a, P_b)$  is achieved by comparing process's structure matrix. The directed graph  $G$  can be represented using an adjacency matrix. To facilitate comparison, graph  $G$  is normalized through the following actions: 1) remove all units of both process that not in  $M$  also with related sequences; 2) add all units from  $M$  that not in one process but in another; 3) any two units of any matching in  $M$  are as equivalent in above steps. For example as is shown in Fig. 2,  $NM_a$  and  $NM_b$  is normalized structure matrix of  $P_a$  and  $P_b$ . In  $NM_a$ ,  $U_1, U_2, U_3, U_4, U_5$  are from  $P_a$  and  $U_6$  is from  $P_b$ ,  $U_2, U_3, U_4$  from both process form a matching in  $M$  and also of  $U_3$  and  $U_4$ . Basing on normalized structure matrix, structure comparison matrix  $CM_{ab}$  is achieved from Eq. 5, where  $ne_{ij}^a$  is element of  $NM_a$ , with position of  $i$ th row and  $j$ th column, and  $ce_{ij}^{ab}$  is element of  $CM_{ab}$ . Finally,  $sim_S(P_a, P_b)$  is calculated to measure the similarity of  $NM_a$  and  $NM_b$  as is shown in Eq. 6, where  $N$  is the number of units in  $CM_{ab}$  and  $x_{ij}$  equals 1 only when  $ce_{ij}^{ab}$  is not -1, otherwise  $x_{ij}=0$ .

	$U_1$	$U_2$	$U_3$	$U_4$	$U_5$	$U_6$
$U_1$	0	1	0	0	0	0
$U_2$	0	0	1	1	0	0
$U_3$	0	0	0	0	1	0
$U_4$	0	0	0	0	1	0
$U_5$	0	0	0	0	0	0
$U_6$	0	0	0	0	0	0

$NM_a$

	$U_1$	$U_2$	$U_3$	$U_4$	$U_5$	$U_6$
$U_1$	0	0	0	0	0	0
$U_2$	0	0	1	1	0	0
$U_3$	0	0	0	0	1	0
$U_4$	0	0	0	0	1	0
$U_5$	0	0	0	0	0	1
$U_6$	0	0	0	0	0	0

$NM_b$

	$U_1$	$U_2$	$U_3$	$U_4$	$U_5$	$U_6$
$U_1$	-1	1	-1	-1	-1	-1
$U_2$	-1	-1	0	0	-1	-1
$U_3$	-1	-1	-1	-1	0	-1
$U_4$	-1	-1	-1	-1	0	-1
$U_5$	-1	-1	-1	-1	-1	1
$U_6$	-1	-1	-1	-1	-1	-1

$CM_{ab}$

Fig. 2 Process Structure Comparison

$$ce_{ij}^{ab} = ne_{ij}^a \wedge ne_{ij}^b = \begin{cases} -1 & , ne_{ij}^a + ne_{ij}^b = 0 \\ 1 & , ne_{ij}^a = 0 \vee ne_{ij}^b = 0 \\ 0 & , ne_{ij}^a = 0 \wedge ne_{ij}^b = 1 \end{cases} \quad (5)$$

$$sim_S(P_a, P_b) = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N x_{ij} ce_{ij}^{ab}}{\sum_{i=1}^N \sum_{j=1}^N x_{ij}} \quad (6)$$

## Application and Discussion

This method was used in building process platform of satellite product. Plate is a kind of structure in satellite and there are many kind of plate such as single layer plate, stiffened plate and sandwich plate which are commonly used now. Although the satellite plates are not completely the same, the architecture and structure of the plates are basic similar to each other, therefore the manufacturing processes of the plates also have high similarity. In order to obtain series of typical process and to make the manufacturing process of plate more generalization and standardization, similarity was measured on process instances so that clustering analysis methods can be applied. Generic process structure of satellite plate achieved by clustering basing on similarity is as shown in Fig. 3.

Before measuring similarity of process, process information need to be structured and readable for computer, while not all of the information in CAPP system are reasonable organization and open for other system, especially those system focus mainly on process card editing. And due to lack of standardization for process information expression and processing can also affect the accuracy of the method in this paper.

In addition, the calculation method of unit similarity, the comparison of proposed approach to other similar one, and possible improvement need to be further studied.

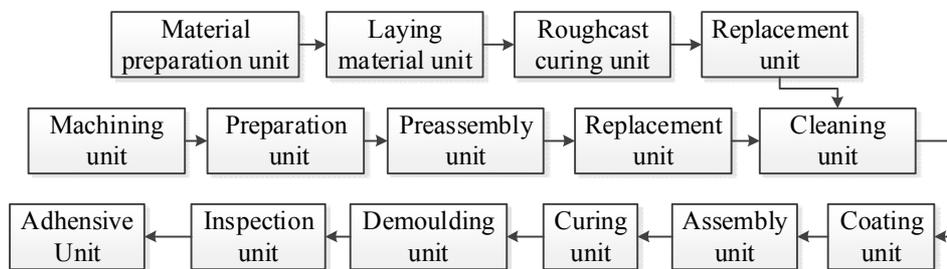


Fig. 3 Generic Process Structure of Satellite Plate

## Acknowledgment

We are grateful to the reviewers and the editors for their constructive suggestions. This work was supported by the National Natural Science Foundation of China (Grant No. 51405104) and the Research Innovation Fund of Harbin Institute of Technology (No. IDGA18102049).

## References

- [1] Jiao, J.R., T.W. Simpson, and Z. Siddique, Product family design and platform-based product development: a state-of-the-art review. *Journal of intelligent Manufacturing*, 2007. 18(1): p. 5-29.
- [2] Zhang, L.L. and R.J. Jiao, A graph rewriting system for process platform planning. *Decision Support Systems*, 2013. 54(2): p. 1174-1191.
- [3] Jiao, J., L. Zhang, and S. Pokharel, Process platform planning for variety coordination from design to production in mass customization manufacturing. *Engineering Management, IEEE Transactions on*, 2007. 54(1): p. 112-129.
- [4] Wang, L., X.-y. Fu, and S.-s. Zhong. Generalized Process Family Modeling Based on Process Platform. in *Proceedings of 20th International Conference on Industrial Engineering and Engineering Management*. 2013. Springer.
- [5] Avramenko, Y. and A. Kraslawski, Similarity concept for case-based design in process engineering. *Computers & Chemical Engineering*, 2006. 30(3): p. 548-557.
- [6] Becker, M. and R. Laue, A comparative survey of business process similarity measures. *Computers in Industry*, 2012. 63(2): p. 148-167.