Machine Component Clustering with Mixing Technique of DSM, Jaccard Distance Coefficient and k-Means Algorithm

Tanongsak Kongsin and Sakon Klongboonjit

January 2, 2020

Machine Component Clustering with Mixing Technique of DSM, Jaccard Distance Coefficient and k-Means Algorithm

Tanongsak Kongsin
Industrial Engineering Department,
Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
e-mail: tanongsak9797@gmail.com

Sakon Klongboonjit
Industrial Engineering Department,
Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
e-mail: sklongboonjit@hotmail.com

***Abstract*—This study aims to introduce a design method for machine component clustering into independent modules so that a machine can be easily modified to achieve its requirement functions. In this study, three techniques of DSM, Jaccard Distance Coefficient and k-Means algorithm are together applied with the 40-component autonomous machine to group all machine components into modules. Clustering steps consist of three steps: 1) Generate a relation matrix, 2) Calculate relationship distance coefficients to build the tree dendrogram and 3) Analyze relationship distance coefficients to find the proper level coefficient. The result shows that the modules of the second level are the most natural.**

Keywords- Modular Design; Jaccard Method; Complete Linkage; k-means algorithm

I. INTRODUCTION

Quick changes in customers' requirements are the essence of today's manufacturing environment. Many companies have a need to change machine functions under the changing production conditions. For flexible manufacturing process enables manufacturers to meet the needs of customers better with a flexible manufacturing system (FMS). [1]. Current methods for flexibility in machine design, designers continue to depend essentially on private decisions [2]. Therefore the modular design of the machine allows changing certain functions to be effective and independent manufacturing [3, 4]. The module construction of the machine requires a relationship between machine components to get a module suitable for change. How to create relationships with parts is achieving through the Design Structure Matrix (DSM) [5] that can show relationships between components. Then use the Complete Linkage method [6] to create a dendrogram with component relationships. This shows the number of modulators generated from the dependency coefficient levels of the six levels [7]. The number of modules used to find the appropriate dependency coefficient using the k-mean algorithm [8, 9] in finding the six levels. With finds from the position of each component of each dependence coefficient to get the appropriate clustering [10].

In this paper, an application of the Complete Linkage method with creating a Dendrogram and k-mean algorithm is introduced to be a new idea to design a modular machine that can easily change machine functions of manufacturing systems.

II. PURPOSE

The main aim of this study is to use clustering with dependency coefficients level and the k-means algorithm to find the dependency coefficient. The main expected outcome is to group the components of the machine into independent modules to make it easier to change machine functions that make the production process flexible and responsive to customer requirements.

III. STUDY METHODOLOGY

In this study, the dependency coefficient and k algorithm are used to find the clustering from the dependency coefficient level and lead to independent module acquisition. To get the expected results, the study process consists of 5 steps as shown in Figure 1.

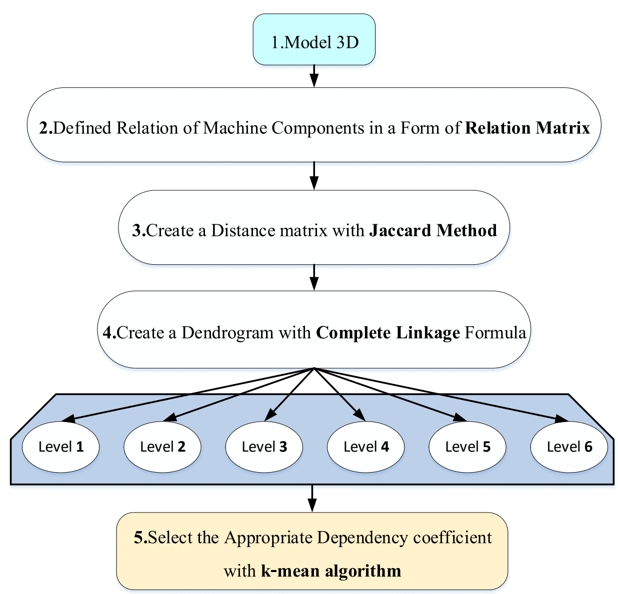

Figure 1. Flow of the proposed method.

IV. RESULTS

A. Model 3D

The automated conveyor system is a 3D model created by designers with a total of 40 parts, as shown in Figure 2.

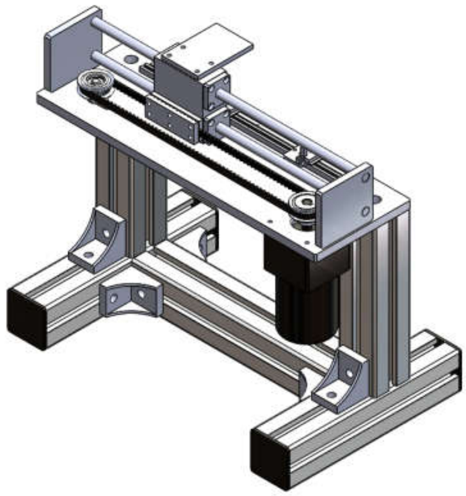

Figure 2. Automated Conveyor System.

B. Defined Relation of Machine Components in a Form of Relation Matrix

Determining the relationship between the components of the machine from the CAD software, which is the main tool for checking the relationship of parts in a 3D model with the following conditions: Relations = 1 means connections and relations = 0 means disconnection, as shown in Figure 3.

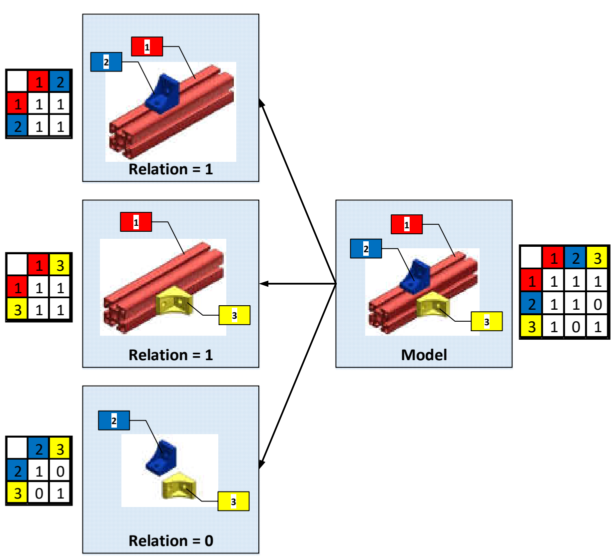

Figure 3. Relationship conditions between parts.

The relationship of each component and its connections components is determined and shown in Figure 4 in the form of Design Structure Matrix (DSM)

NO.	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	…,n
C1	1	1	1								
C2	1	1									
C3	1		1	1							
C4			1	1	1						
C5				1	1						
C6						1					
C7							1	1			
C8							1	1			
C9									1		
C10										1	
…,n											

Figure 4. Example: Design Structure Matrix (DSM)

C. Create a Distance matrix with Jaccard Method

The Jaccard Index formula (Equations 1) is applied to the Design Structure Matrix (DSM) in Figure 4. Finally, the correlation matrix must be transferred to the distance matrix as shown in Figure 5.

$$Jaccard'sim(A,B) = P(A \cap B)P(A \cup B) \quad (1)$$

NO.	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	…,n
C1	0.00	0.67	0.64	0.93	1.00	1.00	1.00	1.00	1.00	1.00	
C2	0.67	0.00	0.90	1.00	0.86	1.00	1.00	0.88	1.00	0.89	
C3	0.64	0.90	0.00	0.64	0.90	1.00	1.00	1.00	1.00	1.00	
C4	0.93	1.00	0.64	0.00	0.67	1.00	1.00	1.00	1.00	1.00	
C5	1.00	0.86	0.90	0.67	0.00	1.00	1.00	0.88	1.00	0.89	
C6	1.00	1.00	1.00	1.00	1.00	0.00	0.83	1.00	1.00	1.00	
C7	1.00	1.00	1.00	1.00	1.00	0.83	0.00	0.50	1.00	1.00	
C8	1.00	0.88	1.00	1.00	0.88	1.00	0.50	0.00	1.00	0.90	
C9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	
C10	1.00	0.89	1.00	1.00	0.89	1.00	1.00	0.90	1.00	0.00	
…,n											

Figure 5. Example: Distance Matrix

D. Create a Dendrogram with Complete Linkage Formula

In this step, the complete linking formula (Equation 2) is used with the distance matrix in Figure 5 to create the dendrogram of this machine, as shown in Figure 6.

$$D((C_i \cup C_j), C_k) = \max_{x \notin c_i, y \notin c_j} \{d(x_i, x_k), d(x_j, x_k)\} \quad (2)$$

The results show the number of clusters at each level of dependency coefficient as follows:

level 1 (Coefficient level = 2.02, Number of clusters = 2.)
level 2 (Coefficient level = 1.95, Number of clusters = 3.)
level 3 (Coefficient level = 1.89, Number of clusters = 4.)
level 4 (Coefficient level = 1.83, Number of clusters = 5.)
level 5 (Coefficient level = 1.75, Number of clusters = 6.)
level 6 (Coefficient level = 1.66, Number of clusters = 9.)

The values in each level are evaluated from Figure 6. by assigning values to be in the range of straight lines as shown in Figure 6.

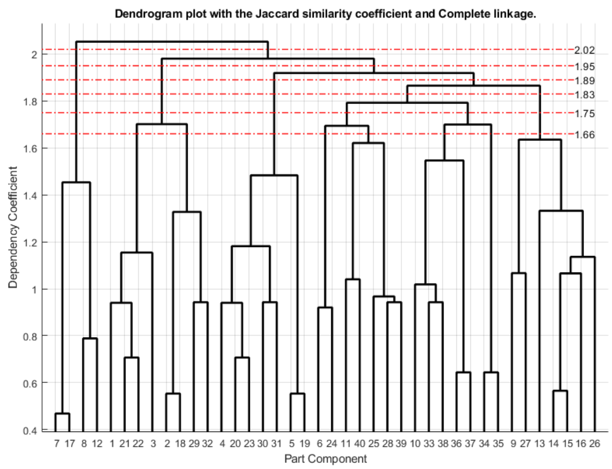

Figure 6. Dendrogram of this machine.

E. Select the Appropriate Dependency coefficient with the k-mean algorithm

Find the position of all 40 component in the x, y, and z-axis, referring to the distance from the machine's center of gravity as shown in Figure 7

$$(\overline{X}\sum V = \sum \overline{X}V, \overline{Y}\sum V = \sum \overline{Y}V, \overline{Z}\sum V = \sum \overline{Z}V) \quad (3)$$

Parts NO.	x	y	z				
1	-180.568	-117.677	-2.039	21	-141.221	-117.777	12.069
2	-180.568	7.323	-27.039	22	-141.221	-117.777	-66.147
3	-15.568	-117.677	-27.039	23	110.324	-117.777	-66.386
4	149.432	-117.677	-2.039	24	129.432	53.364	44.551
5	149.432	7.323	-27.039	25	-67.282	152.299	7.938
6	129.354	131.922	44.56	26	-37.372	132.323	7.961
7	-157.824	134.323	44.051	27	-37.372	172.323	7.961
8	-157.824	121.245	44.051	28	-65.534	189.823	-15.436
9	-15.568	172.323	7.961	29	-180.568	-117.677	113.953
10	-15.568	119.336	-26.15	30	149.432	-117.677	113.953
11	-65.534	134.323	29.993	31	149.432	-117.677	-118.031
12	-157.824	118.323	44.051	32	-180.568	-117.677	-118.031
13	-16.954	112.323	12.571	33	50.074	120.573	-22.439
14	-199.568	149.864	33.943	34	55.515	124.823	-22.439
15	168.432	149.864	33.943	35	65.381	130.816	-22.439
16	-15.568	132.323	7.961	36	-56.259	124.823	-22.439
17	-157.825	137.673	44.615	37	-46.393	130.816	-22.439
18	-180.468	-78.33	12.069	38	-61.7	120.573	-22.439
19	149.332	-78.33	12.069	39	-65.534	146.597	-15.219
20	110.085	-117.777	12.069	40	-14.615	134.323	42.374

Figure 7. x, y, and z-axis coordinates

From the number of clusters that have been imported into the k-means method to select the number of clusters that give the most natural values, select the number of machine modules by the natural values obtained in each cluster as shown in Figures 9

$$\mu_i = \frac{1}{N_i} \sum_{x_1 \in s_1} x_j \quad (4)$$

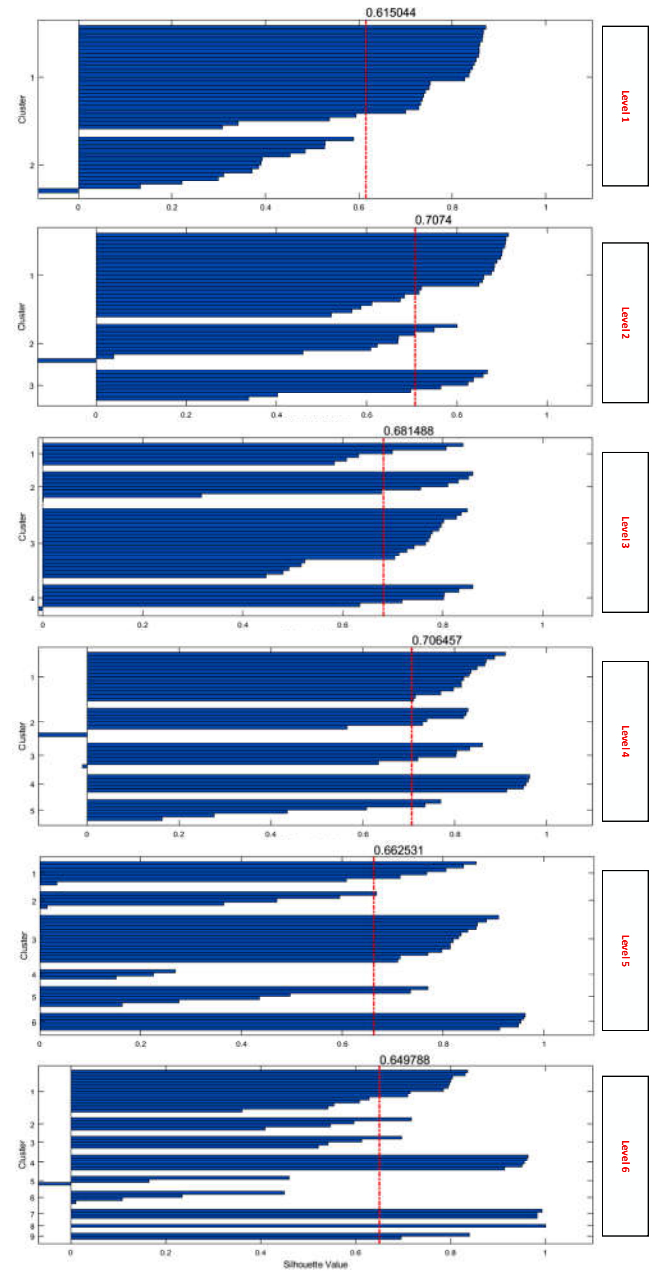

Figure 8. Natural values for clusters 2,3,4,5,6 and 9

Figure 8 shows that the natural value at level 2 has the highest value of 0.7074 and the number of clusters 3

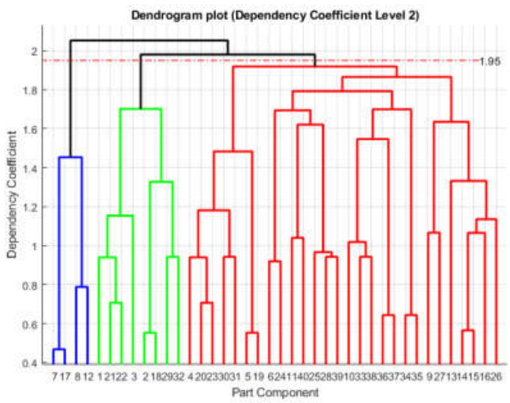

Figure 9. Dendrogram of cutting line at level 2

Figure 9 shows the result of finding the appropriate dependency coefficient level with the k-mean algorithm. The result is 1.949.

Level	Dependency Coefficient	Number of modules	Modules
2	1.949	1	
		2	
		3	

Figure 10. Show the module obtained by cutting at level 2.

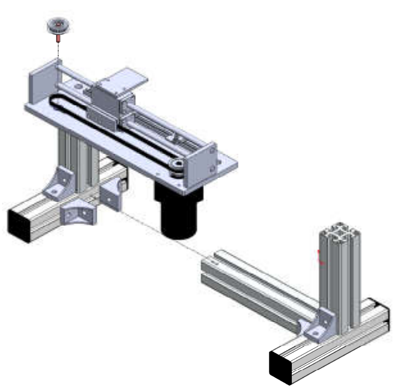

Figure 11. Show of all 3 modules exploded.

Figure 11 shows the result of the module explosion in 3 modules, which can be seen that the module components are appropriate and natural.

V. CONCLUSION

In this article, we studied how to cluster components for automated conveyor systems. In which uses the dependency coefficient level and k- means for clustering analysis. The result is dependency coefficient level = 1.949 and the machine module = 3 the cluster module with 3 modules is considered the most suitable cluster for this automated conveyor system using How to find the dependency coefficient level with k-mean algorithm.

REFERENCES

[1] Shivanand, H., *Flexible manufacturing system*. 2006: New Age International.

[2] Francalanza, E., M. Mercieca, and A. Fenech, *Modular System Design Approach for Cyber Physical Production Systems.* Procedia CIRP, 2018. **72**: p. 486-491.

[3] Bi, Z. and W. Zhang, *Modularity technology in manufacturing: taxonomy and issues.* The International Journal of Advanced Manufacturing Technology, 2001. **18**(5): p. 381-390.

[4] Gauss, L., D.P. Lacerda, and M.A. Sellitto, *Module-based machinery design: a method to support the design of modular machine families for reconfigurable manufacturing systems.* The International Journal of Advanced Manufacturing Technology, 2019: p. 1-26.

[5] Gong, X., Y. Liu, and R.J. Jiao, *Variety-Driven Assembly System Layout Design by Design Structure Matrix Clustering Analysis.* Procedia CIRP, 2017. **63**: p. 295-300.

[6] Kopp, B., *Hierarchical Classification II: Complete - Linkage Method.* Biometrical Journal, 1978. **20**(6): p. 597-602.

[7] Kongsin, T. and S. Klongboonjit. *Machine components clustering with DSM and repeating method: case study of a soil mixing machine*. in *MATEC Web of Conferences*. 2018. EDP Sciences.

[8] Erisoglu, M., N. Calis, and S. Sakallioglu, *A new algorithm for initial cluster centers in k-means algorithm.* Pattern Recognition Letters, 2011. **32**(14): p. 1701-1705.

[9] Jothi, R., S.K. Mohanty, and A. Ojha, *DK-means: a deterministic k-means clustering algorithm for gene expression analysis.* Pattern Analysis and Applications, 2019. **22**(2): p. 649-667.

[10] Llobell, F., et al., *A new approach for the analysis of data and the clustering of subjects in a CATA experiment.* Food quality and preference, 2019. **72**: p. 31-39.