



## Machine Component Clustering with Connection Correlation Method

Tanongsak Kongsin<sup>1</sup>      Sakon Klongboonjit<sup>1\*</sup>

<sup>1</sup>*Industrial Engineering Department, School of Engineering,  
King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand*

\* Corresponding author's Email: sklongboonjit@hotmail.com

---

**Abstract:** This study aims to introduce a new method for modular design, which is called Connection Correlation Method (CCM), to clustering machine parts with inter-part connection conditions to classify independent modules for the Industrial Printer Powertrain set with 105 components. With the CCM technique, module members would be appropriately assigned for each module so that machine modules are independent of their function. The CCM technique should first calculate the distance coefficient matrix with surface contact conditions of bonded, rough, separation, and frictionless for machine components. Finally, this distance coefficient matrix is used to generate the machine dendrogram with the dependency coefficient between 2.4945 and 3.5593. At the dependency coefficient of 2.7913, the 105 components of the Industrial Printer Powertrain set are clustered into 7 modules: Module 1 with 75 interconnection surfaces and 8 parts, Module 2 with 136 interconnection surfaces and 14 parts, Module 3 with 176 interconnection surfaces and 17 parts, Module 4 with 61 interconnection surfaces and 8 parts, Module 5 with 200 interconnection surfaces and 21 parts, Module 6 with 209 interconnection surfaces and 19 parts, and Module 7 with 18 interconnection surfaces and 18 parts. The results show that the CCM technique can apply to design a modular machine like DSM technique, and multitudinous connectivity factors can also be analysed together with general factors.

**Keywords:** Modular design, Connection correlation method (CCM), Inter-part connection conditions, Distance coefficient matrix.

---

### 1. Introduction

Nowadays, industrial machines should adapt themselves along with flexible production processes that machine operators change. To handle with these constraints, modular machines are appropriate to use in flexible production processes. However, it is not easy to design a modular machine with many components; one of the best design methods applied to designing processes is the Machine Modular Design Method [1, 2].

With Machine Modular Design Method, machine designers can design complex modular machines base on the clustering analysis technique, which uses Design Structure Matrix (DSM) to manage design factors [3]. There are now three conventional methods to manage these design factors: 1) Using relationship between component contact surfaces, 2) Using the function of assembling components, and 3)

using the taxonomy of system element interactions [4]. Although these three conventional methods can be applied to design general propose modular machines with independent modules, all conventional methods consider each machine component relationship based on contact or no contact with its neighbor components. That means some contact conditions are overlooked [5], such as contact with sliding, no contact with moving, etc. If a machine designer would like to consider the other contact conditions [6, 7], the multitudinous connectivity factors that can explain more contact conditions between two contact surfaces should be introduced into design processes.

In this study, the relationship between the component multitudinous connectivity factors that can show two more contact conditions of machine component contact surfaces is considered together with conventional contact surfaces. This method is called Connection Correlation Method (CCM), and it

has been introduced to module machine components of the Industrial Printer Powertrain set.

## 2. Purpose

This study aims to introduce the Connection Correlation Method (CCM) for clustering machine components into machine modules. That can be disassembled and easily changed machine functions without affecting the main structure of the machine.

## 3. Study methodology

There are three steps of 1) Connection Correlation Method (CCM), 2) Distance between pairs of observations, and 3) Computing the distance between clusters to cluster machine components into appropriate machine modules with each step has been analyzed sequentially in order. The result of this methodology is the connection correlation value with each step has the detail as following:

### 3.1 Generate connection correlation matrix with connection correlation method (CCM)

The Connection Correlation Method (CCM) is used the explain contact conditions of machine components. This study shows 4 interconnection types of surface contacts: Bonded, Rough, Separation, and Frictionless. The details of each interconnection types are as following:

#### 3.1.1. Bonded

In Fig. 1, Part B is a fixed part, and part A is bonded with part B. This contact condition can explain with the variable  $y_{11}$ , where  $y_{11}$  is a value of 1 when part B is a fixed part, and part A is bonded with part B, and part A can not move; otherwise,  $y_{11}$  is a value of 0.

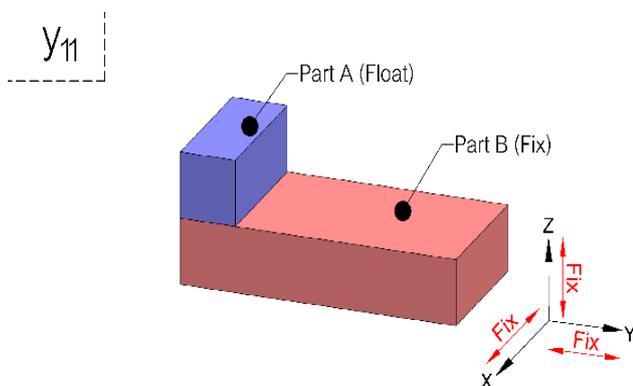


Figure. 1 Condition in the form of variable  $y_{11}$

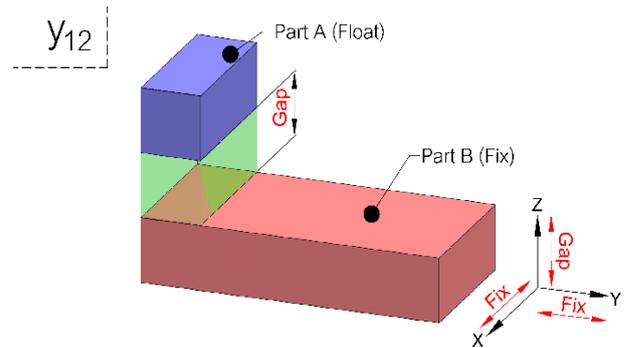


Figure. 2 Condition in the form of variable  $y_{12}$

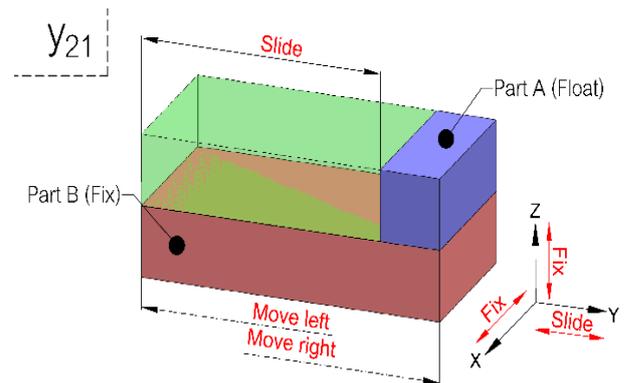


Figure. 3 Condition in the form of variable  $y_{21}$

#### 3.1.2. Rough

In Fig. 2, Part B is a fixed part, and part A is under the rough condition with part B. With the rough condition, the contact surface of part A and part B has a gap. This contact condition can explain with the variable  $y_{12}$ , where  $y_{12}$  is a value of 1 when the contact surface is under the rough condition; otherwise,  $y_{12}$  is a value of 0.

#### 3.1.3. Separation

In Fig. 3, Part B is a fixed part, and part A is under the separation condition with part B. That means part A make contact with part B, and part A can move on their contact surface without friction. This contact condition can explain with the variable  $y_{21}$ , where  $y_{21}$  is a value of 1 when part A can slide on part B without friction; otherwise,  $y_{21}$  is a value of 0.

#### 3.1.4. Frictionless

In Fig. 4, Part B is a fixed part, and part A is under the frictionless condition with part B. With the frictionless condition, the contact surface of part A and part B has a gap, and part A can freely move. This contact condition can explain with the variable  $y_{22}$ , where  $y_{22}$  is a value of 1 when part A is under

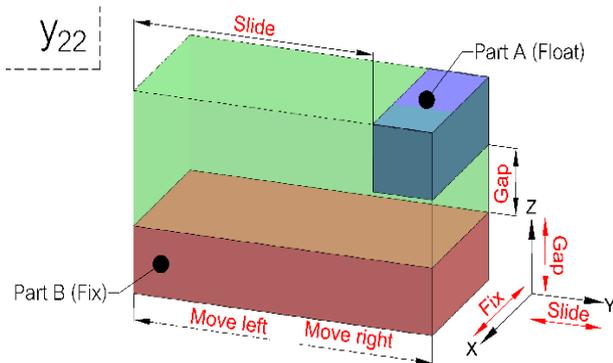


Figure. 4 Condition in the form of variable  $y_{22}$

the frictionless condition; otherwise,  $y_{22}$  is a value of 0.

For any contact surface of two parts, the contact surface can be a mixing condition of bonded, rough, separation, or frictionless. Thus, any two-parts contact surface condition can be shown with  $y_{11}$ ,  $y_{12}$ ,  $y_{21}$ , and  $y_{22}$  like in Eq. (1).

$$\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = y_{ij} \begin{cases} 1, \text{Meet the conditions} \\ 0, \text{Otherwise} \end{cases} \quad (1)$$

where

$y_{11}$  is the connection of bonded.

$y_{12}$  is the connection of rough.

$y_{21}$  is the connection of separation.

$y_{22}$  is the connection of frictionless.

For example, the machine parts in Fig. 5 consists of 4 components. After applying Eq. (1) to explain the contact conditions of all 4 components, the interconnection matrix is generated, as shown in Table 1.

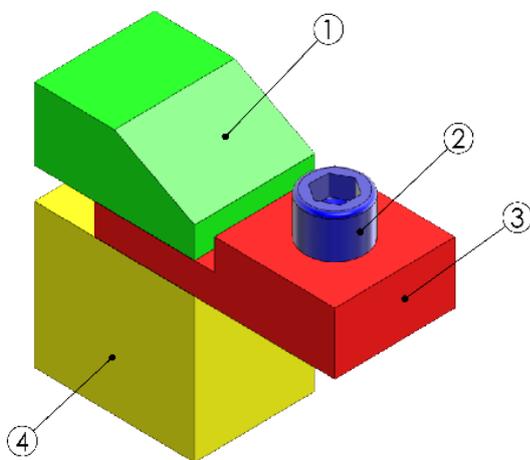


Figure. 5 The model used for describing the method

Table 1. The interconnection matrix condition of all 4 components of the machine part in Fig. 5

No.	1	2	3	4
1	1 1	0 0	1 0	0 0
2	0 0	1 1	1 0	0 0
3	1 0	1 0	1 1	1 0
4	0 0	0 0	1 0	1 1

The values in any cell in Table 1 are the results that analyze by CCM technique. Four values in each cell (From Eq. (1)) show the relationship of any two contact parts, and these four values can be converted into  $CCM_{mn}$  with Eq. (2).

$$CCM_{mn} = \left( \frac{\sum_{i=2}^2 \sum_{j=1}^2 y_{ij}}{4} \right)_{mn} \quad (2)$$

where

$CCM_{mn}$  is the value of the connection correlation, it can be 0 to 1.

For the example, component 1 and 3 in the cell 1,3 has two types of contact surface: bonded ( $y_{11}=1$ ) and separation ( $y_{21}=1$ ) without rough ( $y_{12}=0$ ) and frictionless ( $y_{22}=0$ ). Then, the  $CCM_{13}$  is  $\frac{y_{11}+y_{12}+y_{21}+y_{22}}{4} = \frac{1+0+1+0}{4} = 0.5$  with Eq. (2). After calculation  $CCM_{mn}$  for all calls, the table of values of the connection correlation matrix is generated as shown in Table 2.

From Table 2, component 3 plays the central component role, and the connection correlation values with component 1, component 2, and

Table 2. The connection correlation matrix of all 4 components of the machine part in Fig. 5

No.	1	2	3	4
1	1.00	0.00	0.50	0.00
2	0.00	1.00	0.50	0.00
3	0.50	0.50	1.00	0.25
4	0.00	0.00	0.25	1.00

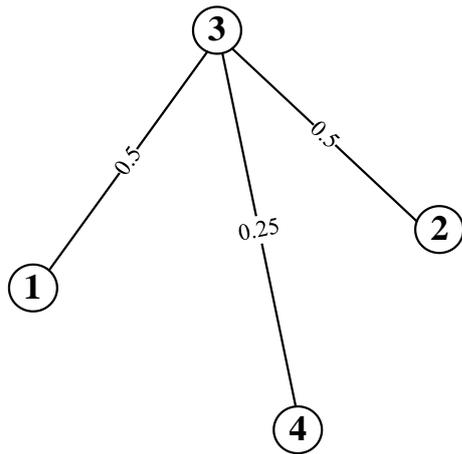


Figure. 6 The connection correlation network of the machine part in Fig. 5

component 4 are 0.5, 0.5, and 0.25, respectively. The connection correlation network of 4 components is shown as in Fig. 6.

### 3.2 Calculate distance between pairs of components

In this step, the distance values of every two nearby components are calculated by Eq. (3), (4), and (5) using Spearman Distance Technique [8]. Use the connection correlation values in Table 3 to generate the machine component distance coefficient matrix in Table 3.

$$d_{st} = 1 - \frac{\sum(r_s - \bar{r}_s)(r_t - \bar{r}_t)}{\sqrt{\sum(r_s - \bar{r}_s)^2} \sqrt{\sum(r_t - \bar{r}_t)^2}} \quad (3)$$

$$\bar{r}_s = \frac{1}{n} \sum_j r_{sj} = \frac{(n+1)}{2} \quad (4)$$

$$\bar{r}_t = \frac{1}{n} \sum_j r_{tj} = \frac{(n+1)}{2} \quad (5)$$

where

$r_{sj}$  is the rank of  $x_{sj}$  taken over  $x_{1j}, x_{2j} \dots x_{mj}$

$r_s$  and  $r_t$  are the coordinate-wise rank vectors of  $x_s$  and  $x_t$

The machine component distance coefficient in Table 3 can show the dissimilarity of any two components. The more machine component distance coefficient is, the more dissimilarity of the two components is. Finally, the machine part dissimilarity network in Fig. 5 is generated as shown in Fig. 7.

Table 3. The matrix of machine component distance coefficient

Part No.	1	2	3	4
1	0.00	1.39	0.50	1.39
2	1.39	0.00	0.50	1.39
3	0.50	0.50	0.00	1.33
4	1.39	1.39	1.33	0.00

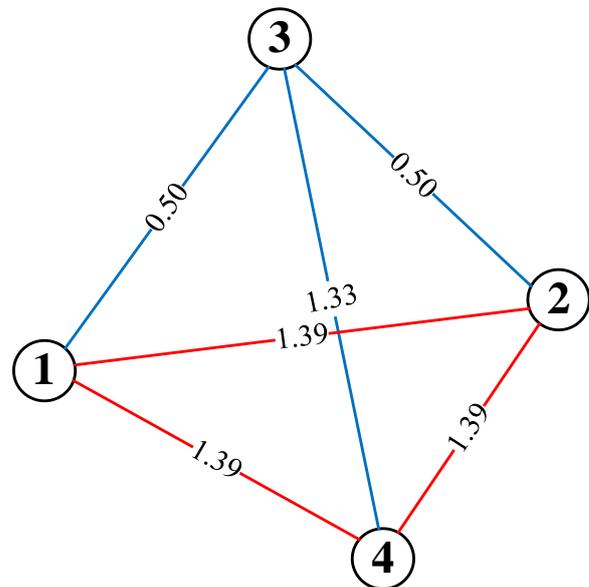


Figure. 7 The dissimilarity network of the machine part in Fig. 5

### 3.3 Generate dendrogram of the machine component distance coefficient

In this step, the dendrogram of the machine component distance coefficient is going to be generated base on the complete-link clustering method with Eq. (6) [9, 10].

$$D(C_k, C_i \cup C_j) = \max \{D(C_k, C_i), D(C_k, C_j)\} \quad (6)$$

where

$D(\cdot, \cdot)$  is a distance between two clusters.

With the complete-link clustering methods, three Dependency Coefficient of 1.1372, 1.9642, and 2.2662 are calculated. Then, the cut-off level values of  $(1.9642 + 2.2662) / 2 = 2.1152$  and  $(1.1372 +$

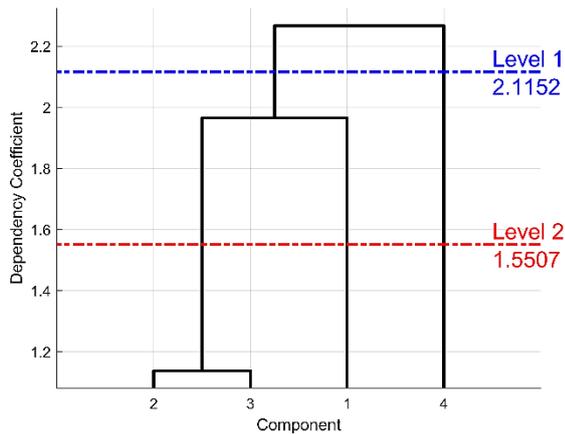


Figure. 8 Dependence coefficient and appropriate clustering methods

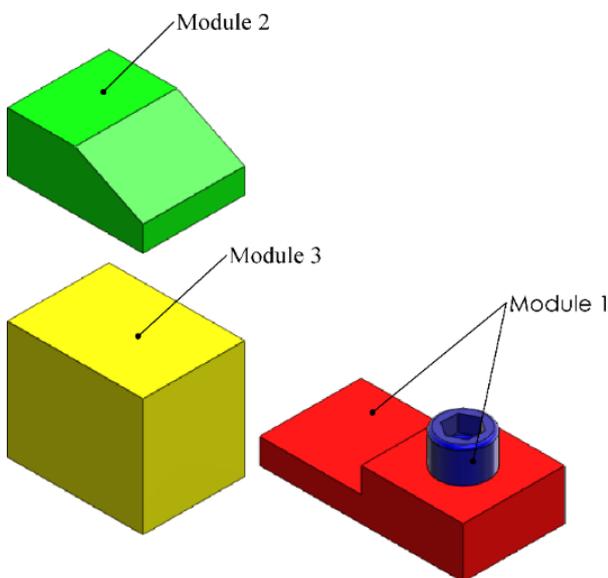


Figure. 9 The 3 modules of the machine part in Fig. 5 at the cut-off level value of 2.1152

$1.9642) / 2 = 1.5507$  are used to clustering the number of modules for the machine part in Fig. 5. The dendrogram of the machine component distance coefficient for the machine part in Figure 5 is created with 2 cut-off level values [11, 12] as shown in Fig. 8.

Finally, the appropriate cut-off level value of 2.1152 is selected to cluster the machine part in Fig. 5 into 3 modules as shown in Fig. 9.

#### 4. Results and discussion

In this study, the industrial printer powertrain set of 105 components in Fig. 10 is being clustered with 3 steps of 1) Generate Connection Correlation Matrix with Connection Correlation Method (CCM), 2) Calculate distance between pairs of components, and 3) Generate dendrogram of the machine component distance coefficient.

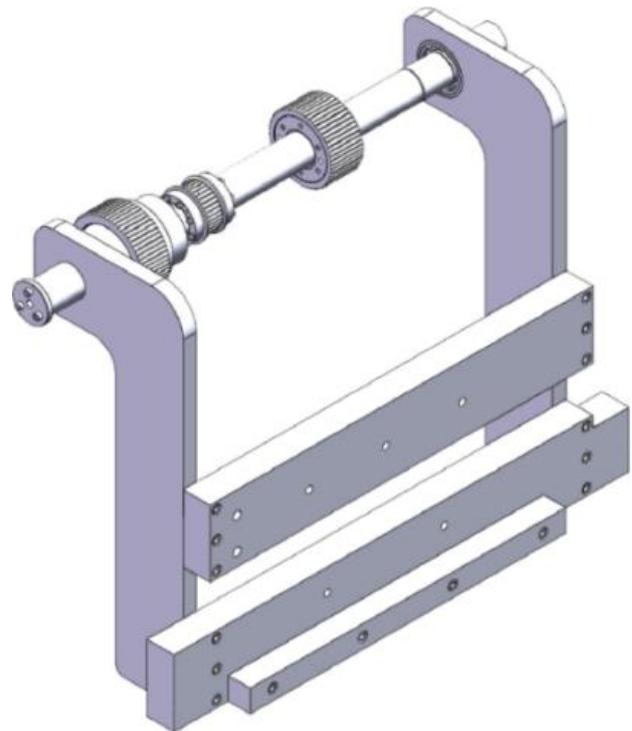


Figure. 10 Industrial Printer Powertrain set

For the first step of Generate Connection Correlation Matrix with Connection Correlation Method (CCM), the interconnection matrix condition of all 105 components is generated with Eq. (1) as shown in Table 5 (matrix size of 105 x 105). With results in Table 4, each cell shows 4 types of interconnection condition for a component  $i$  and  $j$ , the conventional DSM relationships of Bonded and Rough are shown together with the other two relationships of Separation and Frictionless.

Four values in each cell in the interconnection matrix of Table 4 are calculated with Eq. (2) to

Table 4. The interconnection matrix condition of all 105 components of the machine part in Fig. 10

No.	1	2	3	...	105
1	1 1	0 0	1 0	...	0 0
	1 1	0 0	1 0		0 0
2	0 0	1 1	1 0	...	0 0
	0 0	1 1	1 0		0 0
3	1 0	1 0	1 1	...	0 0
	1 0	1 0	1 1		0 0
105	0 0	0 0	0 0	...	1 1
	0 0	0 0	0 0		1 1

Table 5. The connection correlation matrix of all 105 components of the machine part in Fig. 10

No.	1	2	3	...	105
1	1.00	0.00	0.50	...	0.00
2	0.00	1.00	0.50	...	0.00
3	0.50	0.50	1.00	...	0.00
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
105	0.00	0.00	0.00	...	1.00

Table 6. The matrix of machine component distance coefficient of all 105 components of the machine part in Fig. 10

No.	1	2	3	...	105
1	0	0.85	0.61	...	1.05
2	0.85	0	0.61	...	1.05
3	0.61	0.61	0	...	1.06
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
105	1.05	1.05	1.06	...	0

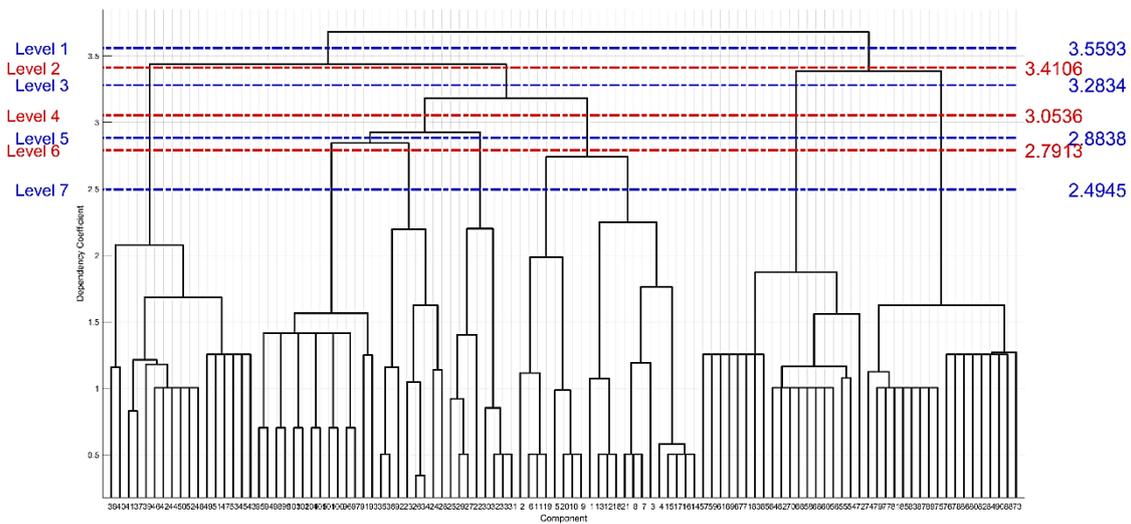


Figure. 11 The dendrogram of all 105 components of industrial printer powertrain set in Fig. 10

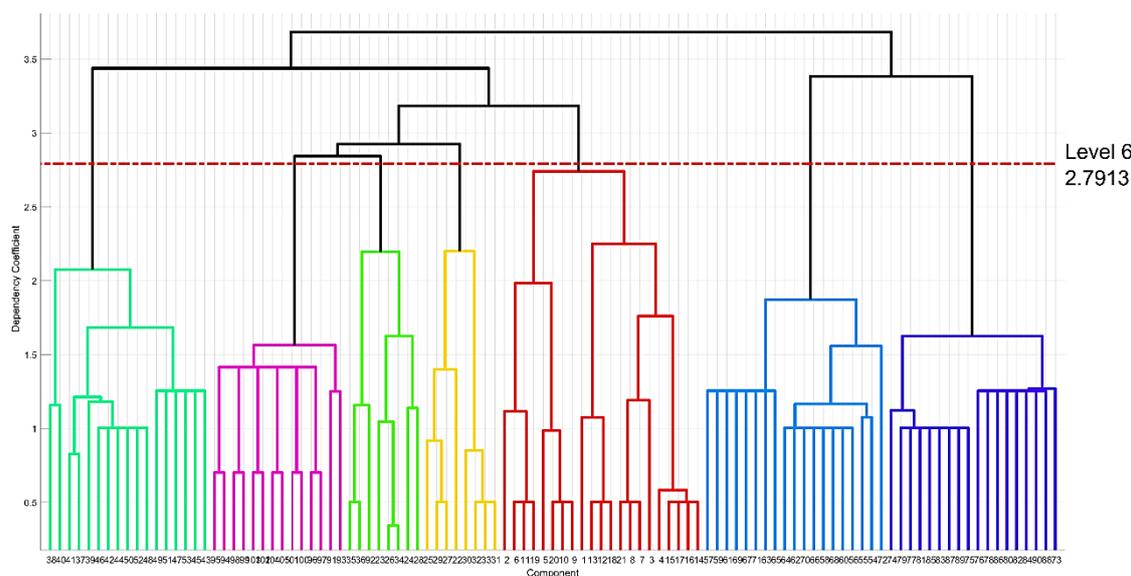


Figure. 12 The dendrogram of all 105 components of industrial printer powertrain set in Fig. 10 at dependency coefficient 2.7913

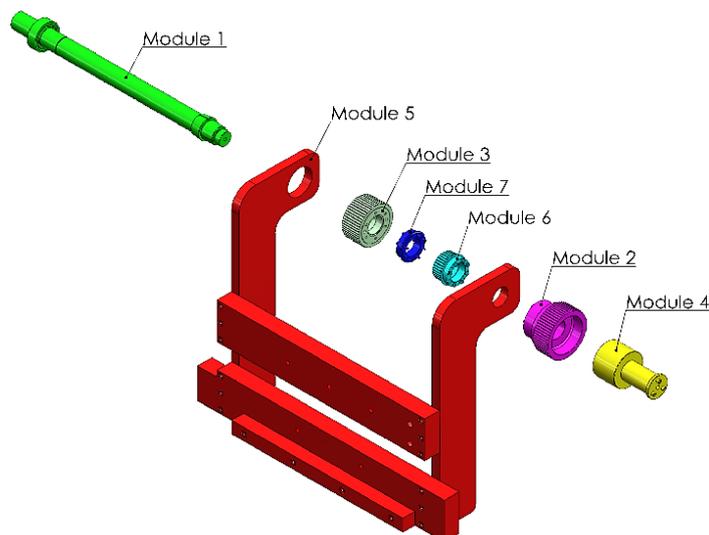


Figure. 13 The 7 modules of industrial printer powertrain set in Fig. 10 at dependency coefficient 2.7913

generate the connection correlation matrix of all 105 components as shown in Table 5. Finally, the connection correlation matrix of all 105 components is complete and ready to transfer into the machine component distance coefficient matrix in the next step.

For the second step of Calculate distance between pairs of components, the Spearman Distance Technique with Eq. (3) is applied to all values in Table 5 to generate the matrix of machine component distance coefficient of all 105 components of the machine part in Fig. 10 as shown in Table 6.

For the last step of Generate dendrogram of the machine component distance coefficient, the complete-link clustering method in Eq. (6) is applied to Table 6 to generate the dendrogram of all 105 components of the machine part in Fig. 10 with a 7-level Dependency Coefficient of 3.5593 (Level 1), 3.4106 (Level 2), 3.2834 (Level 3), 3.0536 (Level 4), 2.8838 (Level 5), 2.7913 (Level 6), and 2.4945 (Level 7) as shown in Fig. 11.

Finally, the appropriate cut-off level value of 2.7913 is selected to cluster the machine part in Fig. 10 into 7 modules as shown in Fig. 12 (The dendrogram form) and in Fig. 13 (The module form).

The 1<sup>st</sup> module is with 75 interconnection surfaces and 8 pieces of the machine part. The 2<sup>nd</sup> module is with 136 interconnection surfaces and 14 pieces of the machine part. The 3<sup>rd</sup> module is with 176 interconnection surfaces and 17 pieces of the machine part. The 4<sup>th</sup> module is with 61 interconnection surfaces and 8 pieces of the machine part. The 5<sup>th</sup> module is with 200 interconnection surfaces and 21 pieces of the machine part. The 6<sup>th</sup> module is with 209 interconnection surfaces and 19 pieces of the machine part. The 7<sup>th</sup> module is with 18

interconnection surfaces and 18 pieces of the machine part.

## 5. Conclusion

This study introduces the new method (CCM) to analyze all machine component interconnection surface relationships with the conventional DSM relationships cooperate with other relationships [3]. With CCM, every interconnection surface relationship of machine component will not be overlooked, unlike the DSM method only with the binary interconnection surface relationships. With the example of the Industrial Printer Powertrain set, the conventional DSM relationships of Bonded and Rough can be considered together with the other two relationships of Separation and Frictionless. The clustering results should undoubtedly differ from the conventional DSM method. For further study, other than the general surface relationship of machine components such as the physical surface relationship, the heat transfer surface relationship, the mechanical surface relationship, etc., can be input during machine designers' machine design processes.

## Conflicts of Interest

The authors declare no conflict of interest.

## Author Contributions

Conceptualization, Tanongsak Kongsin and Sakon Klongboonjit; methodology, Tanongsak Kongsin; software, Tanongsak Kongsin; validation, Tanongsak Kongsin and Sakon Klongboonjit; formal analysis, Sakon Klongboonjit; investigation, Tanongsak Kongsin; resources, Sakon Klongboonjit; data curation, Sakon Klongboonjit; writing—original

draft preparation, Tanongsak Kongsin; writing—review and editing, Tanongsak Kongsin; visualization, Tanongsak Kongsin; supervision, Sakon Klongboonjit; project administration, Tanongsak Kongsin; funding acquisition, Sakon Klongboonjit.

## References

- [1] E. Greve, C. Rennpferdt, and D. Krause, “Harmonizing cross-departmental perspectives on modular product families”, *Procedia CIRP*, Vol. 91, No. 4, pp. 452-457, 2020.
- [2] C. Krahe, A. Bräunche, A. Jacob, N. Stricker, and G. Lanza, “Deep Learning for Automated Product Design”, *Procedia CIRP*, Vol. 91, No. 4, pp. 3-8, 2020.
- [3] T. Kongsin and S. Klongboonjit, “Machine Component Clustering with Mixing Technique of DSM, Jaccard Distance Coefficient and k-Means Algorithm”, In: *Proc. of International Conference on Industrial Engineering and Applications (ICIEA)*, Bangkok, Thailand, pp. 251-255, 2020.
- [4] Y. Ogasawara and M. Kon, “Two clustering methods based on the Ward's method and dendrograms with interval-valued dissimilarities for interval-valued data”, *International Journal of Approximate Reasoning*, Vol. 129, No. 1, pp. 103-121, 2021.
- [5] P. Kumar and A. Kanavalli, “A Similarity based K-Means Clustering Technique for Categorical Data in Data Mining Application”, *International Journal of Intelligent Engineering and Systems*, Vol. 14, No. 2, pp. 43-51, 2021.
- [6] D. H. Eggink and M. W. Groll, “Joining element design and product variety in manufacturing industries”, *Procedia CIRP*, Vol. 88, No. 13, pp. 76-81, 2020.
- [7] X. Geng and H. Tang, “Clustering by connection center evolution”, *Pattern Recognition*, Vol. 98, No. 1, pp. 107063-107075, 2020.
- [8] A. Kumar and S. Abirami, “Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method”, *Information Sciences*, Vol. 460, No. 1, pp. 23-41, 2018.
- [9] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*, A John Wiley & Sons, United States of America, 2005.
- [10] A. K. Abasi, A. T. Khader, M. A. Al-Betar, S. Naim, S. N. Makhadmeh, and Z. A. A. Alyasseri, “Link-based multi-verse optimizer for text documents clustering”, *Applied Soft Computing*, Vol. 87, No. 1, pp. 106002-106025, 2020.
- [11] C. Tadonki, F. Meyer, and F. Irigoien, “Dendrogram based algorithm for dominated graph flooding”, *Procedia Computer Science*, Vol. 29, No. 1, pp. 586-598, 2014.
- [12] M. Azhar, J. Z. Huang, M. A. Masud, M. J. Li, and L. Cui, “A hierarchical Gamma Mixture Model-based method for estimating the number of clusters in complex data”, *Applied Soft Computing*, Vol. 87, No. 1, pp. 105891-105910, 2020.