Zhang et al.
Chinese Journal of Mechanical Engineering (2023) 36:17
https://doi.org/10.1186/s10033-023-00841-5

Chinese Journal of Mechanical Engineering

ORIGINAL ARTICLE

Open Access

Product Specification Analysis for Modular Product Design Using Big Sales Data

Jian Zhang[1], Bingbing Li[1], Qingjin Peng[2] and Peihua Gu[3,4*]

Abstract

Big data on product sales are an emerging resource for supporting modular product design to meet diversified customers' requirements of product specification combinations. To better facilitate decision-making of modular product design, correlations among specifications and components originated from customers' conscious and subconscious preferences can be investigated by using big data on product sales. This study proposes a framework and the associated methods for supporting modular product design decisions based on correlation analysis of product specifications and components using big sales data. The correlations of the product specifications are determined by analyzing the collected product sales data. By building the relations between the product components and specifications, a matrix for measuring the correlation among product components is formed for component clustering. Six rules for supporting the decision making of modular product design are proposed based on the frequency analysis of the specification values per component cluster. A case study of electric vehicles illustrates the application of the proposed method.

Keywords Modular product design, Customer preference, Product specifications, Correlation analysis, Big sales data, Electric vehicle

1 Introduction

Innovative product design is essential for global competition among manufacturers because the lifecycle performance of a product is highly influenced by the quality of product design solutions [1, 2]. Compared with the technology-driven mode of product design based on manufacturers' technical abilities, the market-driven mode of product design can rapidly respond to market changes with the required product specifications [3–5].

*Correspondence:
Peihua Gu
peihua.gu@tju.edu.cn
[1] Department of Mechanical Engineering, Shantou University, Shantou 515063, China
[2] Department of Mechanical Engineering, University of Manitoba, Manitoba R3T 5V6, Canada
[3] International Institute for Innovative Design and Intelligent Manufacturing of Tianjin University, Shaoxing 312000, China
[4] School of Mechanical Engineering, Tianjin University, Tianjin 300350, China

The required changes in product specifications and their combinations can be achieved through adaptations of the physical components. In the past decades, various types of product structures have been proposed to develop modular products to satisfy the changing market requirements, such as mass manufactured products, mass customized products, reconfigurable products, upgradeable products, open architecture products, and adaptable products [6–8]. A comparison of the different types of products is presented in Table 1 [8]. It was observed that these modular products consist of the following modules and interfaces.

- Common module: Common modules are designed and manufactured by manufacturers using mass production methods [9, 10]. Those shared by different products are typically considered platforms in a product family.
- Customized module: Customized modules are designed and manufactured by the original equip-

Table 1 Types of products with different modules

Type of product	Definition	Example	Included modules and interfaces			
			Common module	Customized module	Personalized module	Adaptable interface
Mass produced product	Product obtained through large-scale production for meeting standard requirements	Standardized bearings	Yes	No	No	No
Mass customized product	Product designed and manufactured to meet customized requirements with approximate efforts of large-scale production	Electric vehicles that provide customized features for user selections	Yes	Yes	No	No
Reconfigurable product	Product that can meet different requirements through rapid reconfigurations of components can be reconfigured to meet different requirements	Cameras can be reconfigured with different lenses to meet various photographing requirements	Yes	Yes	No	Yes
Upgradeable product	Product that can accommodate new requirements or technology advancement by upgrading or replacing components	Computers that can be upgraded by replacing their memory modules	Yes	Yes	No	Yes
Open architecture product	Product with a platform, such that add-on modules developed by different vendors can be connected to the platform through interfaces [7, 12]	Excavators that can be connected to different front execution devices provided by third-party vendors	Yes	Yes	Yes	Yes
Adaptable product	Product that can be easily adapted in the operation stage to meet different requirements [6]	Blades of a wind turbine can be adapted to meet different operation conditions	Yes	Yes	Yes	Yes

ment manufacturer (OEM) using mass customization methods for customer choice during the purchasing process [11].

- Personalized module: Personalized modules are designed and manufactured to satisfy the personalized requirements of the customers. They can be designed by customers and/or purchased from a third-party vendor.
- Adaptable interface: Adaptable interfaces are components and their relationships to the interface can be easily adapted in the product operation stage to satisfy different requirements [12, 13].

To better satisfy the changing customer/market requirements of product specifications, there is a need to support the design decisions of common, customized, and personalized modules and their interfaces. In the past decades, many research efforts have been devoted to modular design methods (such as module-based product family design, platform design, and modular design for mass customization) for better product variety, reduced production costs, and shortened lead time. Different approaches, such as the design structure matrix method and modular function deployment method, have been proposed for product modularization. Various metrics (such as the component shape, materials, and manufacturing capability) for modularity evaluation have been proposed for modular product design [14]. To identify common modules shared by a product family, platform and product family design have been proposed by researchers. Methods for customized module determination have also been developed in modular design for mass customization.

Despite progress in modular product design, it was observed that most of the existing methods focused on identifying common (platform) and customized modules. Few studies have been conducted on the identification and design of personalized modules and adaptable interfaces for connecting personalized modules. In addition, existing methods tend to use independence and/or similarity based on considerations of component geometry, materials, assembly, disassembly, etc. for product modularity. Unambiguous functional correlation analysis among components remains a less regarded topic in product modularization [15]. It was observed that the prediction and incorporation of diversified market preferences on product specification combinations have rarely been studied to support modular design decisions.

Customers and markets have different preferences for product specification combinations. Design decisions on different types of modules (i.e., common, customized, personalized modules) and their interfaces to satisfy various customer requirements need to be carried out based on market preference analysis. The correlations between specifications can effectively reflect customer and market preferences for product specification combinations. Generally, relationships among product specifications may have three categories: Independent, normal, and strong. These relationships are illustrated in Figure 1.

To better facilitate the design of different types of modular products to satisfy diversified changeable requirements, specification correlations originating from customer/market preferences should be investigated for design decision-making of physical components by considering specification-component relationships [16]. However, there is a lack of research on the correlations between product specifications for product modularity [17, 18]. Although existing methods for customer preference analysis (such as individual judgment, expert meeting, brainstorming, Delphi method, and Kano method) have been widely used, owing to their strong applicability and good interpretability, the following limitations have been observed:

- The dataset used is not large enough because only a limited number of customers can participate in the interview, and only parts of the customers provide useful feedback or comments on the product. This results in an insufficient reflection of customer preferences in the marketplace.
- Product specification correlations originating from both conscious and subconscious customer preferences were not fully considered. Because most existing methods focus on analyzing customer preferences for certain functions and/or features of a product, few studies have investigated specification correlations originating from customer preferences and their combinations.

Big data on product sales are valuable sources for mining relationships among product specifications. Specification relations originating from customers' conscious and subconscious preferences are embedded in big data on product sales. Related research has demonstrated that sales data can provide critical guidance for product design and development [16, 19]. This study proposes a method for the correlation analysis of product specifications and components using big sales data to support modular product design. In this study, the sales data of target products are collected and used to reflect customer preferences. Specification correlations were obtained according to the sales and specification values. Additionally, four types of relations between specifications and components were defined so that designers could categorize these relations according to different design requirements. According to the correlations of the product components, different combinations of components were

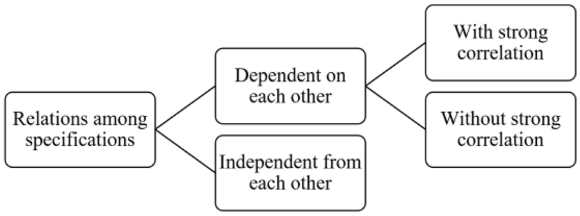

Figure 1 Relationships among specifications

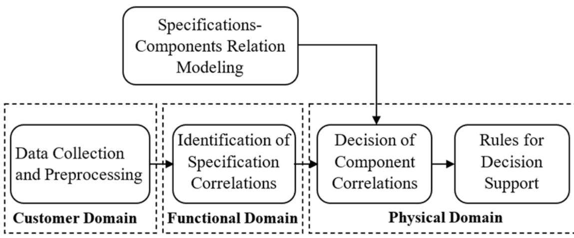

Figure 2 Framework of the proposed method

obtained using a hierarchical clustering algorithm. Component combinations are used to support design decision making for the updating/development of products.

2 Proposed Method

Product specifications in the functional domain were designed to meet customer requirements. The physical components are designed to support the specifications required by customers. Correlations among product specifications are the basis of component design and reconfiguration. According to statistical analysis based on the collected product data, a method for measuring correlations among product specifications is proposed based on the following assumptions:

(1) Product sales data, such as the number of product sales and combinations of product specifications, can be obtained through market surveys.
(2) Correlations among product specifications can originate from the customers' conscious and subconscious preferences.
(3) The correlations among product specifications originating from customer preferences are embedded in the product sales data.

Based on these three assumptions, a framework is proposed for modular product design decision support through the quantification of correlations among product specifications, as shown in Figure 2.

The method associated with this framework is illustrated in Figure 3. The proposed method comprises five phases. In Phase 1, the sales data, product specifications, and their values are obtained through a market survey, and these data are processed to support the search for specification correlations. In Phase 2, a specification correlation matrix is formed to measure the correlations among product specifications. The matrix reflects customer preferences when searching using sales data. In Phase 3, four relations between product components and specifications are identified to satisfy different design requirements. Based on the relations between specifications and components, the correlations of product components are determined in Phase 4. Using the component correlations, the components are clustered into different groups using a hierarchical clustering method in Phase 5. Additionally, a frequency analysis of the corresponding specification values per component cluster is conducted to obtain the characteristics of each cluster. Finally, rules for supporting the decision making of modular product design are formed based on the component clustering results.

2.1 Data Collecting and Preprocessing

In this step, a market survey is conducted to acquire relevant data for each product instance. The market survey was implemented by accessing the official websites of these products and other database platforms. The survey output included various products and their sales data, specifications, and values.

Following data collection, the data were preprocessed. Screening of non-relevant specifications can be performed manually to meet the scope and objectives of the investigation. Consequently, the redundant specifications can be removed. Furthermore, a dimensionality reduction operation may be performed by removing entries in product instances and specifications to reduce the large amount of useless data [20, 21].

In this study, supposed m product instances are collected, and the sales number of products is modelled as follows:

$$\boldsymbol{N} = [N_1, \ldots, N_m], \tag{1}$$

where N_i represents the sales of the ith product and i =1, 2, ..., m.

Within the collected datasets, it is assumed that n specifications are considered for the correlation analysis, and the specifications can be modelled as follows:

$$\boldsymbol{S} = [S_1, \ldots, S_n], \tag{2}$$

where S_i (i=1, ..., n) denotes the ith specification.

To simplify the calculation, the values of each product specification need to be digitized/discretized in the data preparation for correlation searching. For example, for

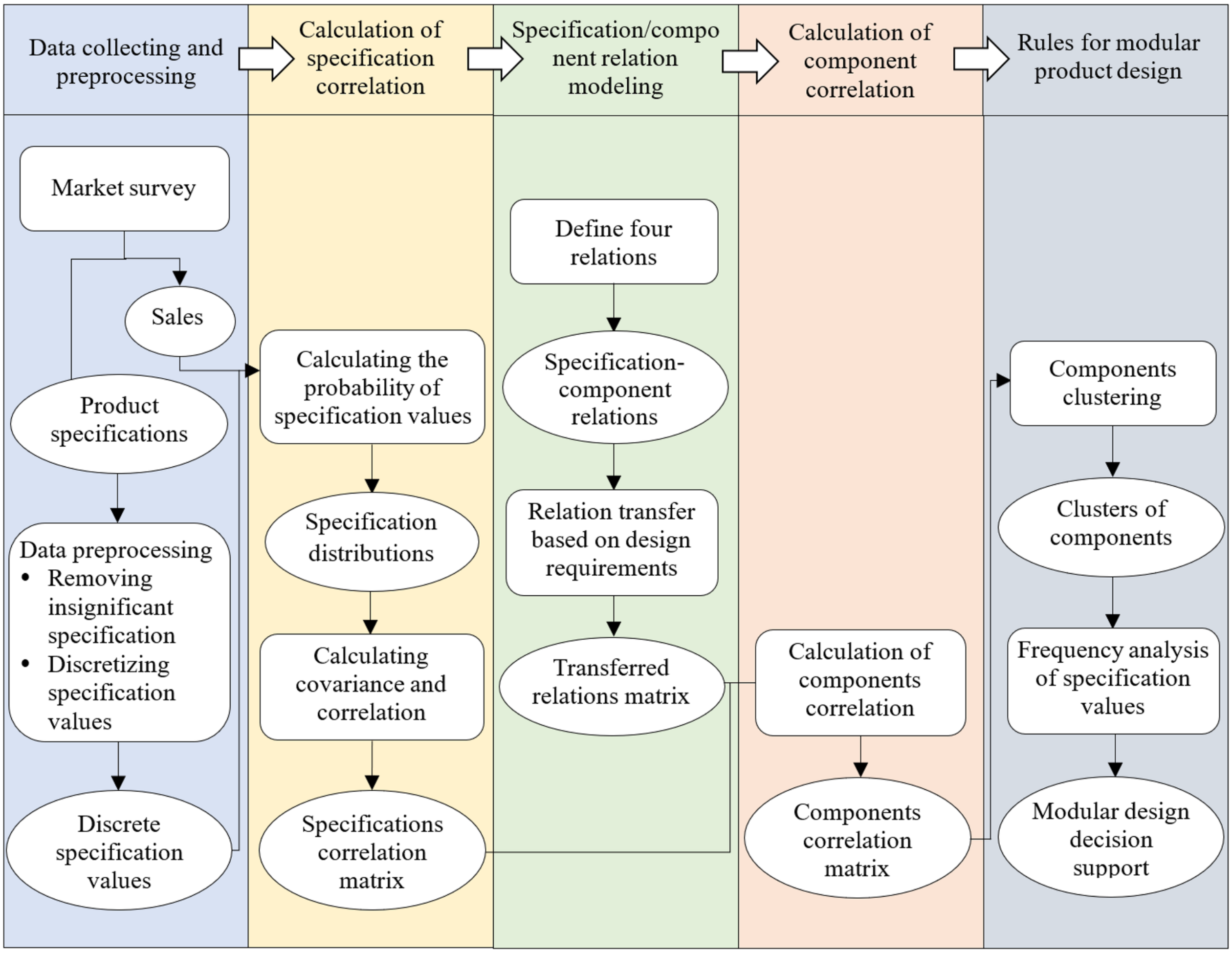

Figure 3 Five phases of the proposed method

battery cell shapes (S_i), a value of 1 or 2 can be assigned to represent the shape of a cylinder or square, respectively.

2.2 Correlations among Product Specifications

In this step, correlations among product specifications are calculated using the collected big sales data. Generally, correlations between two specifications originating from customers' conscious and subconscious preferences can be linear or nonlinear.

The correlation between two variables can be described using correlation coefficients. The Pearson correlation coefficient [22] and Spearman rank correlation coefficient [23] are commonly used to measure the linear and nonlinear relationships between two variables, respectively. The values of the correlation coefficient varied between −1 and 1.

In this study, correlation coefficients are required to measure correlations between specifications, considering customer preferences through big data on product sales. The correlations among all the n specifications can be obtained using the following matrix $\boldsymbol{M_S}$:

$$\boldsymbol{M_S} = \begin{bmatrix} 1 & \rho_{S_1S_2} & \cdots & \rho_{S_1S_n} \\ \rho_{S_1S_2} & 1 & \cdots & \rho_{S_2S_n} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{S_1S_n} & \rho_{S_2S_n} & \cdots & 1 \end{bmatrix}, \quad (3)$$

where $\rho_{S_iS_j}$ is the correlation between S_i and S_j, $i, j = 1, 2, \ldots, n$. n is the number of product specifications.

2.3 Specifications and Components Relations

Product specifications are physically realized by the components. The relationships between the components and specifications of a product are required for the investigation of component correlations.

In logic and mathematics, necessary and sufficient are used to describe the conditional relationship between the two statements. If the truth of statement A guarantees the truth of statement B, then A is sufficient to B, and B is necessary for A. On this basis, four types of relations between product specifications and components are considered, as follows.

(1) Neither sufficient nor necessary. The component is relatively independent of the specification, which means that specification changes do not cause component changes, and component changes do not cause specification changes.
(2) Sufficient and unnecessary. The component is affected by multiple specifications, which means that specification changes will cause a component change, but the component change does not necessarily cause a specification change.
(3) Necessary but not sufficient. The specification is affected by multiple components, which means that specification changes do not cause a component change; however, component changes will cause a specification change.
(4) Sufficient and necessary. The specification only affects the component, and the component is only affected by the specification.

To achieve different design requirements or objectives, the above four types of relations between product specifications and components must be transferred into influential (I) and non-influential (N). The following three scenarios related to the design objective were considered for the relationship transformation:

(a) Scenario 1: The design objective is to minimize the number of changing components while satisfying the specification changes. For example, the modular design of an open architecture product satisfies changeable specification requirements. In this situation, the relations of types 2 and 4 are defined as influential (I), whereas 1 and 3 are non-influential (N) when changes in specifications are known. This scenario is applicable to the condition that the priority target is to meet the changeable requirements of the product specifications.
(b) Scenario 2: The design objective is to minimize the number of specifications that could be influenced by the required adaptations of the components. For example, the modular design of products for upgrading or replacing required components. In this situation, the relations of types 3 and 4 are defined as influential (I), whereas 1 and 2 are non-influential (N) when changes in components are known. This scenario is applicable under the condition that the priority target is to facilitate the required upgrading or adaptations of certain product components.
(c) Scenario 3: The design objective is to maximize the adaptation capability to satisfy unknown changes in both the specifications and components. For example, the modular design of adaptable products satisfies potential changes in both product specifications and components. In this situation, the relations of types 2, 3, and 4 are defined as influential (I) and type 1 as non-influential (N) when changes in the components and specifications are unknown. This scenario is applicable under the condition that potential changes in both specifications and components should be considered.

The relation transfers based on different design objectives and their application conditions are summarized in Table 2.

In this study, a matrix $\boldsymbol{M_R}$ is proposed to model the transferred relations of the product specifications and components as follows:

$$\boldsymbol{M_R} = \left[r_{ij}\right]_{m\times n}, \tag{4}$$

where r_{ij} represents the relationship between the ith specification and jth component. r_{ij} uses I and N to represent the influential and non-influential relationships between the ith specification and jth component, respectively.

2.4 Correlations among Product Components

Component correlations matrix $\boldsymbol{M_C}$ can be formed based on the correlations of specifications $\boldsymbol{M_S}$ and specification-component relations $\boldsymbol{M_R}$. A computational method is proposed to map specification correlations in the customer domain to component correlations in the physical domain. Component correlations can be used for decision making in the design of physical product structures.

Table 2 Relation transfer based on different design objectives

Scenario	Design objectives	Processing	Condition of application
1	Minimize the number of changing components while satisfying the specification changes	2&4→I 1&3→N	Satisfy changeable requirements of product specifications
2	Minimize the number of specifications influenced by required changes of components	3&4→I 1&2→N	Facilitate required component upgrading or adaptation
3	Maximize the adaptation capability to satisfy unknown changes of specifications and components	2&3&4→I 1→N	Satisfy potential changes of both specifications and components

In this study, we assume that t is the number of components considered in the design, and specifications S_p, …, S_q are influenced by components C_x, x=1, …, t, $p \neq q$, p, $q \in [1, n]$. Specifications S_u, …, S_v are influenced by physical components C_y, y=1, …, t, $u \neq v$, u, $v \in [1, n]$.

The correlations between components C_x and C_y can be determined based on the correlations among specifications S_p, …, S_q and S_u, …, S_v. In this study, the correlations between components C_x and C_y are defined by the mean values of the correlations among related specifications S_p, …, S_q and S_u, …, S_v. Therefore, the correlations between the components C_x and C_y can be determined by

$$\rho_{C_xC_y} = \frac{\sum_{j=u}^{v}\sum_{i=p}^{q}\rho_{S_iS_j}}{(q-p)\times(v-u)}, \tag{5}$$

where x, y=1, …, t, $p \neq q$, $u \neq v$, p, q, u, v = 1, 2, …, n; n denotes the number of specifications considered in the design.

According to Eq. (5), the correlation among the components can be modelled by the correlation matrix $\boldsymbol{M_C}$ as follows:

$$\boldsymbol{M_C} = \begin{bmatrix} 1 & \rho_{C_1C_2} & \cdots & \rho_{C_1C_t} \\ \rho_{C_1C_2} & 1 & \cdots & \rho_{C_2C_t} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{C_1C_t} & \rho_{C_2C_t} & \cdots & 1 \end{bmatrix}, \tag{6}$$

Obviously, the matrix has the following properties:

- It is a symmetric matrix, which indicates non-direction in the component correlations.
- The values of the diagonal elements are equal to one, which reflects the correlations between the components themselves.

2.5 Components Clustering and Frequency Analysis

Product components can be clustered to obtain rules and implications for design decisions using the formed matrix of the product component $\boldsymbol{M_C}$. The clustering results formed the component groups.

In this step, a hierarchical clustering method is used to cluster the components. Hierarchical clustering is an unsupervised learning method that supports data clustering without the target attributes. The data were evaluated to determine some intrinsic structures of the data. Hierarchical clustering hierarchically decomposes a given dataset into a tree using two schemes: Cohesive and split. Aggregated hierarchical clustering is a bottom-up approach. Split hierarchical clustering is contrary to agglomerated hierarchical clustering [24, 25].

The result of hierarchical clustering is a set of components as a significant basis for product design reorganization. Components can be designed in the same module when their correlations are high, for a convenient solution for design improvement. For convincing design decision support, clustering results should be analyzed to meet the scope of applications.

In the frequency analysis, a specification group i (Figure 4) can be obtained according to the clustering result and specification/component relations, which is a set of all specifications related to the components in cluster i.

Frequency analysis searches for each specification in cluster i according to the collected data. Kurtosis is used to measure the degree of data aggregation to reflect the distribution of the specification values. The kurtosis of each specification can be obtained through frequency analysis. A large kurtosis value means that the specification values are relatively concentrated, which can be considered a dominant specification. However, a small kurtosis value indicates a high divergence in the specification values. In this sense, medium and small kurtosis values should be considered for customized and personalized modules, respectively.

Using this concept, two thresholds t_1 and t_2 are empirically defined based on the designers' experience, where $t_1 > t_2$. Clusters can then be defined for different types based on the following rules.

(1) If all kurtosis in this cluster is larger than a predefined threshold value t_1, the cluster can be considered a common module.
(2) If the kurtosis in the cluster has a value among the predefined threshold values t_1 and t_2, the cluster can be considered a customized module.
(3) If the kurtosis in the cluster has a value less than the predefined threshold value t_2, the cluster can be considered a personalized module.

2.6 Rules for Supporting Decision-Making of Modular Product Design

A product module forms a unit of components with a particular function [26]. In other words, some elements of a product are combined to realize specific functions.

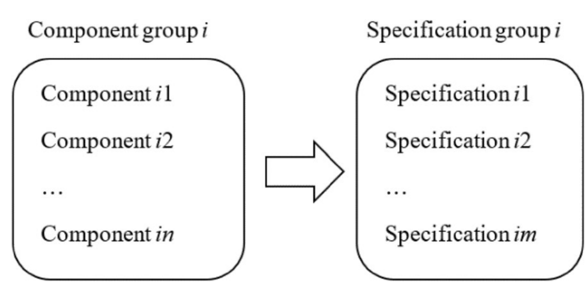

Figure 4 Searching for specification groups

Customized modules can be changed according to customer requirements, but the options are limited. For a personalized module, users can add modules according to their personalized requirements. Open adaptable interfaces allow third-party modules from different sources to meet the customers' changeable requirements [13].

Components with different characteristics should be structured into different modules [27]. Therefore, different design decisions can be made for the components in each cluster according to the kurtosis of all the specifications in different clusters.

In this study, the following rules are proposed to support modular design decisions based on the clustering results.

a) Components in the same cluster are recommended to be grouped into one module.
b) Components in different clusters are recommended not to be grouped into the same module.
c) If all kurtosis in a cluster is larger than a predefined threshold value t_1, the cluster should be formed as a common module.
d) If the kurtosis in the cluster has a value between the predefined threshold values t_1 and t_2, it is suggested that the cluster form a customized module.
e) If the kurtosis in the cluster has a value less than the predefined threshold value t_2, the cluster is suggested to be a personalized module.
f) Adaptable interfaces are suggested to connect customized and personalized modules with the common module.

Considering the diversity of products in the market, the proposed rules can be used to support the design decisions for different types of products as shown in Figure 5.

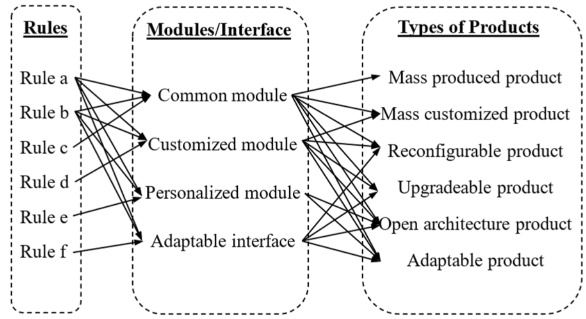

Figure 5 Applications of the rules

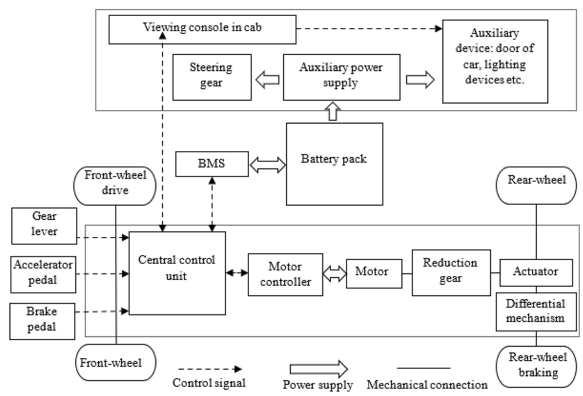

Figure 6 Typical structure of EVs

3 Case Study

Electric vehicles (EVs) have attracted significant attention worldwide in the application of green technology. In recent years, a significant number of EVs has emerged in the market. Although EVs show promise in the automotive market, increasing demand and fast upgrades cause chaotic challenges in the EV market. For design upgrading, the lack of consideration of customer preferences is a concern, and it is particularly important to build upgrading rules of EVs based on customer preferences. In this research, the proposed method is used for the update of EV components to support design decision-making in EV renewal.

A typical structure of an EV is shown in Figure 6, which illustrates the functional modules of the vehicle [28]. To meet diversified customer requirements, EVs should be well designed with adaptability using a modular or open architecture. The design support of EVs based on the correlations among the product specifications is illustrated using the proposed method.

3.1 Data Collection and Preprocess of EVs

Using an online search, 31 different types of EVs were identified. Data including sales and specification values of the EVs were collected. It should be noted that only sales in the Chinese market in 2018 were recorded. The data were obtained from the official websites of these products. Some of the collected datasets are listed in Table 3.

The specifications of EVs were then obtained by evaluating and analyzing the existing EVs in the market. Considering that partial specifications have the same characteristics without details of the data, 15 were selected after data preprocessing for analysis. Additionally, as listed in Table 4, the values of the 15 specifications were discretized.

Table 3 Collected data sets

ID	Sales in 2018	Specifications						
		Price (×1000 CNY)	Electric mileage (km)	Total power (kW)	Total torque (N·m)	Energy density (W·h/kg)	Quick charge time (h)	Battery capacity (kW·h)
1	46213	23.07	400	160	310	160.8	1.5	60.48
2	43634	27.41	350	160	310	140.97	1.5	61.9
3	35699	9.99	305	160	310	146.27	0.5	43.2
4	31426	23.83	400	120	250	135.4	0.5	52
5	27870	17.19	200	41.8	150	140.91	0.5	23.6
6	16102	20.29	318	80	230	122.68	0.5	48
7	15336	5.98	255	30	90	150	0.5	27
8	10329	18.98	401	90	276	144.05	0.5	54.3
9	8852	24.65	410	132	290	212	0.5	54.75
10	7484	14.58	353	120	250	142.07	0.5	52

Table 4 Product specifications and values

ID	Specification	Values	ID	Specification	Values
S_1	Price (×1000 CNY)	[0,10]; (10,20]; (20,30]; (30,40]; (40,+∞);	S_2	Electric mileage (km)	[0,200]; (200,300]; (300,400]; (400,500]; (500,+∞);
S_3	Engine type (horsepower)	[0,90]; (90,120]; (120,150]; (150,180]; (180,210]; (210,+∞);	S_4	Weight (kg)	[0,1000]; (1000,1300]; (1300,1600]; (1600,1900]; (1900,2100]; (2100,+∞);
S_5	Motor power (kW)	[0,50]; (50,100]; (100,150]; (150,200]; (200,+∞);	S_6	Top speed(km/h)	[0,100]; (100,120]; (120,140]; (140,160]; (160,+∞);
S_7	Total motor torque (N·m)	[0,100]; (100,200]; (200,300]; (300,400]; (400,500]; (500,+∞);	S_8	Electricity of 100 km (kW·h)	[0,11]; (11,13]; (13,15]; (15,17]; (17,19]; (19,+∞);
S_9	Energy density(W·h/kg)	[0,135]; (135,145]; (145,155]; (155,165]; (165,+∞);	S_{10}	Quick charge time (h)	[0,0.5]; (0.5,1]; (1,1.5]; (1.5,+∞); No;
S_{11}	Battery capacity (kW·h)	[0,25]; (25,40]; (40,55]; (55,70]; (70,+∞);	S_{12}	Size of central console screen (inch)	[0,5]; (5,7]; (7,9]; (9,11); (11,+∞);
S_{13}	Cell model	Cylindrical; Square;	S_{14}	GPS	Yes; No;
S_{15}	Power sunroof	Yes; No;			

3.2 Calculation of Specification Correlations

In this study, the specification distributions and correlations were obtained according to the collected data described in Section 3.1. Pairwise correlations among the 15 specifications were calculated in a 15×15 matrix $\boldsymbol{M_s}$ as partially shown below for specifications S_1 to S_{15}. As shown in $\boldsymbol{M_s}$, for example, the correlation between S_2 and S_6 is 0.93, which means that the probability of S_2 and S_6 changing together is very high, according to the customer preferences.

3.3 Relation Modeling of Specifications and Components

The major components considered in this design are listed in Table 5. A list of components is formed by analyzing the structures of the sample EVs. The characteristics of each component are then analyzed to identify the relationships between the components and specifications, as shown in Table 6. In this case study, Scenario 2, which is summarized in Table 2, was considered. Therefore, the transferred relation matrix $\boldsymbol{M_R}$ can be obtained from Table 6.

$$\boldsymbol{M_s} = \begin{bmatrix}
1 & 0.40 & 0.74 & 0.45 & 0.27 & 0.48 & 0.48 & 0.38 & 0.01 & 0.18 & 0.61 & 0.18 & 0.27 & 0.13 & 0.19 \\
0.40 & 1 & 0.55 & 0.55 & 0.62 & 0.93 & 0.63 & 0.39 & 0.30 & 0.04 & 0.75 & 0.30 & 0.03 & 0.27 & 0.56 \\
0.74 & 0.55 & 1 & 0.62 & 0.65 & 0.60 & 0.61 & 0.45 & 0.15 & 0.28 & 0.78 & 0.48 & 0.07 & 0.18 & 0.58 \\
0.45 & 0.55 & 0.62 & 1 & 0.77 & 0.64 & 0.88 & 0.78 & 0.08 & 0.10 & 0.83 & 0.29 & 0.34 & 0.12 & 0.49 \\
0.27 & 0.62 & 0.65 & 0.77 & 1 & 0.68 & 0.86 & 0.30 & 0.45 & 0.14 & 0.77 & 0.23 & 0.13 & 0.26 & 0.67 \\
0.48 & 0.93 & 0.60 & 0.64 & 0.68 & 1 & 0.73 & 0.41 & 0.24 & 0.15 & 0.74 & 0.27 & 0.04 & 0.16 & 0.64 \\
0.48 & 0.63 & 0.61 & 0.88 & 0.86 & 0.73 & 1 & 0.58 & 0.26 & 0.08 & 0.84 & 0.23 & 0.27 & 0.16 & 0.52 \\
0.38 & 0.39 & 0.45 & 0.78 & 0.30 & 0.41 & 0.58 & 1 & 0.31 & 0.01 & 0.58 & 0.34 & 0.47 & 0.07 & 0.33 \\
0.01 & 0.30 & 0.15 & 0.08 & 0.45 & 0.24 & 0.26 & 0.31 & 1 & 0.31 & 0.25 & 0.12 & 0.24 & 0.14 & 0.14 \\
0.18 & 0.04 & 0.28 & 0.10 & 0.14 & 0.15 & 0.08 & 0.01 & 0.31 & 1 & 0.38 & 0.20 & 0.12 & 0.15 & 0.15 \\
0.61 & 0.75 & 0.78 & 0.83 & 0.77 & 0.74 & 0.84 & 0.58 & 0.25 & 0.38 & 1 & 0.49 & 0.30 & 0.03 & 0.51 \\
0.18 & 0.30 & 0.48 & 0.29 & 0.23 & 0.27 & 0.23 & 0.34 & 0.12 & 0.20 & 0.49 & 1 & 0.20 & 0.06 & 0.47 \\
0.27 & 0.03 & 0.07 & 0.34 & 0.13 & 0.04 & 0.27 & 0.47 & 0.24 & 0.12 & 0.30 & 0.20 & 1 & 0.04 & 0.27 \\
0.13 & 0.27 & 0.18 & 0.12 & 0.26 & 0.16 & 0.16 & 0.07 & 0.14 & 0.15 & 0.03 & 0.06 & 0.04 & 1 & 0.03 \\
0.19 & 0.56 & 0.58 & 0.49 & 0.67 & 0.64 & 0.52 & 0.33 & 0.14 & 0.15 & 0.51 & 0.47 & 0.27 & 0.03 & 1
\end{bmatrix}. \tag{7}$$

$$\boldsymbol{M_C} = \begin{bmatrix}
1 & 0.61 & 0.62 & 0.62 & 0.62 & 0.67 & 0.67 & 0.67 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 1 & 0.68 & 1 & 1 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.62 & 0.62 \\
0.61 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.66 & 0.66 & 0.65 & 0.64 & 0.64 & 0.64 & 0.64 & 0.61 & 0.57 & 0.61 & 0.61 & 0.61 & 0.64 & 0.64 & 0.64 & 0.64 & 0.67 & 0.67 \\
0.62 & 0.67 & 1 & 1 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.65 & 0.65 & 0.65 & 0.65 & 0.62 & 0.58 & 0.62 & 0.62 & 0.62 & 0.65 & 0.65 & 0.65 & 0.65 & 1 & 1 \\
0.62 & 0.67 & 1 & 1 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.65 & 0.65 & 0.65 & 0.65 & 0.62 & 0.58 & 0.62 & 0.62 & 0.62 & 0.65 & 0.65 & 0.65 & 0.65 & 1 & 1 \\
0.62 & 0.67 & 1 & 1 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.65 & 0.65 & 0.65 & 0.65 & 0.62 & 0.58 & 0.62 & 0.62 & 0.62 & 0.65 & 0.65 & 0.65 & 0.65 & 1 & 1 \\
0.67 & 0.66 & 0.67 & 0.67 & 0.67 & 1 & 1 & 1 & 0.65 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.62 & 0.67 & 0.67 & 0.67 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.67 \\
0.67 & 0.66 & 0.67 & 0.67 & 0.67 & 1 & 1 & 1 & 0.65 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.62 & 0.67 & 0.67 & 0.67 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.67 \\
0.67 & 0.66 & 0.67 & 0.67 & 0.67 & 1 & 1 & 1 & 0.65 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.62 & 0.67 & 0.67 & 0.67 & 0.66 & 0.66 & 0.66 & 0.66 & 0.67 & 0.67 \\
0.63 & 0.65 & 0.66 & 0.66 & 0.66 & 0.65 & 0.65 & 0.65 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.63 & 0.62 & 0.63 & 0.63 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 0.66 & 0.66 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
1 & 0.61 & 0.62 & 0.62 & 0.62 & 0.67 & 0.67 & 0.67 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 1 & 0.68 & 1 & 1 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.62 & 0.62 \\
0.68 & 0.57 & 0.58 & 0.58 & 0.58 & 0.62 & 0.62 & 0.62 & 0.62 & 0.60 & 0.60 & 0.60 & 0.60 & 0.68 & 1 & 0.68 & 0.68 & 0.68 & 0.60 & 0.60 & 0.60 & 0.60 & 0.58 & 0.58 \\
1 & 0.61 & 0.62 & 0.62 & 0.62 & 0.67 & 0.67 & 0.67 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 1 & 0.68 & 1 & 1 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.62 & 0.62 \\
1 & 0.61 & 0.62 & 0.62 & 0.62 & 0.67 & 0.67 & 0.67 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 1 & 0.68 & 1 & 1 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.62 & 0.62 \\
1 & 0.61 & 0.62 & 0.62 & 0.62 & 0.67 & 0.67 & 0.67 & 0.63 & 0.64 & 0.64 & 0.64 & 0.64 & 1 & 0.68 & 1 & 1 & 1 & 0.64 & 0.64 & 0.64 & 0.64 & 0.62 & 0.62 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.64 & 0.64 & 0.65 & 0.65 & 0.65 & 0.66 & 0.66 & 0.66 & 0.64 & 1 & 1 & 1 & 1 & 0.64 & 0.60 & 0.64 & 0.64 & 0.64 & 1 & 1 & 1 & 1 & 0.65 & 0.65 \\
0.62 & 0.67 & 1 & 1 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.65 & 0.65 & 0.65 & 0.65 & 0.62 & 0.58 & 0.62 & 0.62 & 0.62 & 0.65 & 0.65 & 0.65 & 0.65 & 1 & 1 \\
0.62 & 0.67 & 1 & 1 & 1 & 0.67 & 0.67 & 0.67 & 0.66 & 0.65 & 0.65 & 0.65 & 0.65 & 0.62 & 0.58 & 0.62 & 0.62 & 0.62 & 0.65 & 0.65 & 0.65 & 0.65 & 1 & 1
\end{bmatrix}. \tag{8}$$

Table 5 Components of an EV

ID	Component	ID	Component	ID	Component
C_1	Chassis	C_2	Drive motor	C_3	Motor controller
C_4	Central control unit	C_5	BMS	C_6	Gear lever
C_7	Accelerator pedal	C_8	Brake pedal	C_9	Battery pack
C_{10}	Front mounting	C_{11}	Front wheel	C_{12}	Front wheel brake
C_{13}	Steering gear	C_{14}	Car body	C_{15}	Body adornment
C_{16}	Car door	C_{17}	Seat	C_{18}	Body shell
C_{19}	Rear Suspension	C_{20}	Rear wheel	C_{21}	Rear wheel braking
C_{22}	Transmission shaft	C_{23}	Differential mechanism	C_{24}	Reducer

Table 6 Specifications-components relation matrix

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
C_1	3	1	1	3	1	1	1	1	1	1	1	1	1	1	1
C_2	3	3	3	3	4	3	4	1	1	1	1	1	1	1	1
C_3	3	3	1	3	1	3	1	1	1	2	1	1	1	1	1
C_4	3	3	1	3	1	3	1	1	1	2	1	1	1	1	1
C_5	3	3	1	3	1	3	1	1	1	2	1	1	1	1	1
C_6	1	1	1	1	1	3	1	1	1	1	1	1	1	1	1
C_7	1	1	1	1	1	3	1	1	1	1	1	1	1	1	1
C_8	1	1	1	1	1	3	1	1	1	1	1	1	1	1	1
C_9	3	3	1	3	1	3	1	3	2	2	4	1	2	1	1
C_{10}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{11}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{12}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{13}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{14}	3	1	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{15}	3	1	1	1	1	1	1	3	1	1	1	2	1	2	2
C_{16}	3	1	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{17}	3	1	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{18}	3	1	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{19}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{20}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{21}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{22}	3	3	1	3	1	1	1	1	1	1	1	1	1	1	1
C_{23}	3	3	1	3	1	3	1	1	1	1	1	1	1	1	1
C_{24}	3	3	1	3	1	3	1	1	1	1	1	1	1	1	1

3.4 Component Correlations

The component correlations are searched based on the specification correlation matrix $\boldsymbol{M_S}$ and relation matrix $\boldsymbol{M_R}$.

In this case, 24 components were identified, and the correlation matrix of the components $\boldsymbol{M_C}$ (as shown in Eq. (8)) was obtained using Eq. (6). The results of the component correlations provide a basis for component clustering. The values in $\boldsymbol{M_C}$ reflect the probability of the clustered components based on the customer preferences.

3.5 Component Clustering and Frequency Analysis

Hierarchical clustering was conducted based on the correlation matrix, $\boldsymbol{M_C}$. In this process, every item in the matrix is replaced by the difference between the original value and 1.

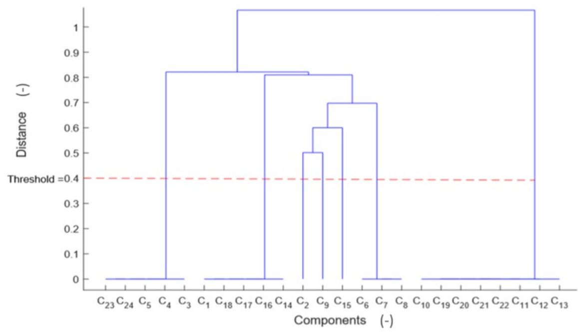

Figure 7 Results of component clustering

As shown in Figure 7, different component clusters can be obtained when different thresholds are selected. Considering the existing product structure, when the threshold is lower than 0.5, the clustering results can satisfy the physical constraints of the product components to the greatest extent. When the threshold is greater than 0.5, C_2 (drive motor) and C_9 (battery pack) are clustered into the same cluster. This is inconsistent with the existing design. Therefore, in this case study, the threshold is set to 0.4. The clustering results and corresponding specification groups are listed in Table 7.

As illustrated in Table 7, high correlations among the components were observed in the same cluster. In Cluster 1, the results showed that C_{10}, C_{11}, C_{12}, C_{13}, C_{19}, C_{20}, C_{21}, and C_{22} should be formed in a module. However, in the existing design of EVs, C_{10}, C_{11}, C_{12} C_{13}, are structured in a module, and C_{19}, C_{20}, C_{21}, and C_{22} are designed in a module. This is because the correlations among the specifications were not properly evaluated in the original design. In cluster 2, C_3, C_4, C_5, C_{23}, and C_{24} can form a module. From Table 7, it can be observed that C_{23} and C_{24} should be separated from this cluster when considering the physical structure. In cluster 3, C_1, C_{14}, C_{16}, C_{17} and C_{18} were physically related. This demonstrates the rationality of the existing design of EVs.

In Figure 7, the distances between the components in clusters 1, 2, 3, and 4 are zero, which means that these

Table 7 Component clusters of EVs

Clusters	Components	Specification groups
Cluster 1	C_{10}, C_{11}, C_{12}, C_{13}, C_{19}, C_{20}, C_{21}, C_{22}	S_1, S_2, S_4
Cluster 2	C_3, C_4, C_5, C_{23}, C_{24}	S_1, S_2, S_4, S_6
Cluster 3	C_1, C_{14}, C_{16}, C_{17}, C_{18}	S_1, S_4
Cluster 4	C_6, C_7, C_8	S_6
Cluster 5	C_2	S_1, S_2, S_4, S_6, S_7
Cluster 6	C_9	S_1, S_2, S_4, S_6, S_8
Cluster 7	C_{15}	S_1, S_8

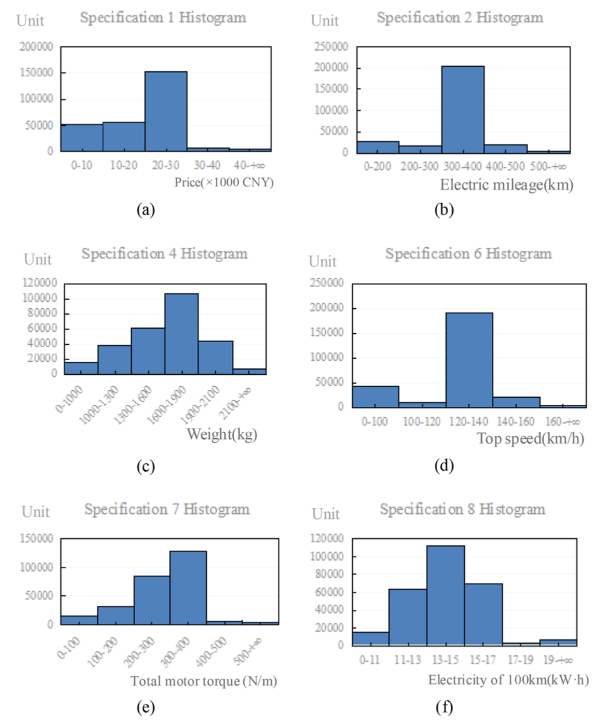

Figure 8 Part of specification frequency analysis

components should not be clustered into different clusters, regardless of the threshold. In industry, sufficient specification data must be collected for product design to improve the accuracy of the proposed method. To better support the design decision, a specification frequency analysis was conducted as shown in Figure 8. In this analysis, the number of samples was determined based on the sales of different types of products.

The kurtosis of the 15 specifications is obtained as shown in Figure 9. For the clusters, if all kurtosis values in a specification group are larger than t_1, the components in this cluster should be designed as a common module. If there is kurtosis in the specification between t_1 and t_2, the components in this cluster can be designed as

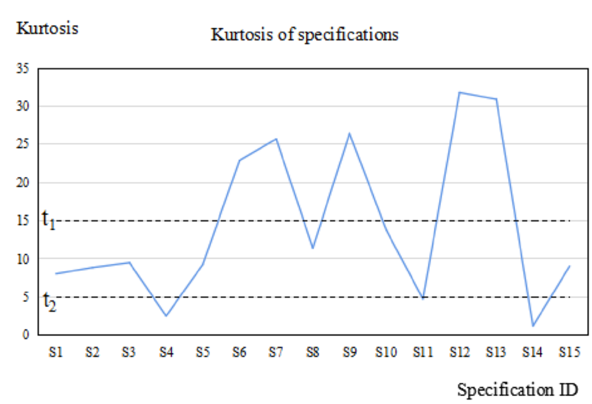

Figure 9 Kurtosis of all specifications

Table 9 Cluster analysis results

Clusters	Components	Types of modules
Cluster 1	C_{10}, C_{11}, C_{12}, C_{13}, C_{19}, C_{20}, C_{21}, C_{22}	Personalized
Cluster 2	C_3, C_4, C_5, C_{23}, C_{24}	Personalized
Cluster 3	C_1, C_{14}, C_{16}, C_{17}, C_{18}	Personalized
Cluster 4	C_6, C_7, C_8	Common
Cluster 5	C_2	Personalized
Cluster 6	C_9	Personalized
Cluster 7	C_{15}	Customized

a customized module. If the kurtosis of the specification is less than t_2, the components in this cluster should be designed as personalized modules.

3.6 Modular Design Decision Support of EVs

According to the cluster analysis results in Section 3.5, the clusters of the components were divided into three categories: Common, customized, and personalized modules as illustrated in Table 9. Additional modular product design recommendations of EVs are presented in Table 10.

From Table 10, it can be observed that recommendations 1, 2, 4, and 6 are consistent with the existing design, which proves the effectiveness of the proposed method. However, it is obvious that recommendation 3 does not fit the existing design of EVs, because C_6, C_7, and C_8 are not designed as common components. Based on the analysis results, C_6, C_7, and C_8 should be redesigned as common components or grouped into a common module. Compared with the existing design of EVs that does not include a personalized module, recommendation 5 suggests that a personalized module should be designed to satisfy the personalized requirements of the vehicle in the marketplace.

It is worth mentioning that the design recommendations summarized in Table 10 may be unsuitable for the design modifications of a specific existing product because the physical constraints of the manufacturer technique capability also need to be considered. Additionally, metrics related to the component geometry, materials, assembly, and disassembly should be considered for product modularity evaluations in real modular product design practices. Integrations of the newly proposed method with existing modular design methods must be carried out for modular product design.

4 Discussions and Conclusions

Increasing market driving and technology development require a rapid response of products to meet market and customer requirements through product specification combinations. Specification relations originating from customer conscious and subconscious preferences can be embedded in the big sales data. To better facilitate various types of modular product designs to satisfy the changeable market/customer requirements, relations among product specifications should be evaluated using big data on product sales. In this study, a framework and associated method were proposed to support modular product design decisions based on the correlation analysis of product specifications and components using the big sales data. Correlations among product specifications were identified by analyzing product sales data. By considering the components and specification relations, a matrix for measuring the correlation among the product components was formed for component cluster analysis. Frequency analysis of the corresponding specification values per component cluster was performed to evaluate the dominance of the component clusters. Six rules for supporting the decision making of modular product design were proposed based on the frequency analysis results. A case study of EVs was used to illustrate the proposed method. The contributions of this study are summarized as follows.

- The newly proposed specification correlation can reflect customer conscious and subconscious prefer-

Table 10 Decision support of modular EV design

ID	Modular product design decision recommendations	Rules
1	Components in the same cluster are recommended to be grouped into the same module (e.g., C_6, C_7, C_8 are recommended to be grouped into the same module)	a
2	Components in different clusters are recommended not to be grouped into one module (e.g., C_2 and C_9 are not recommended to be grouped into the same module)	b
3	Components in cluster 4 are recommended to be designed as common components or modules. (e.g., C_6, C_7, and C_8 are recommended to be designed as common components or being grouped into a common module)	c
4	Values of S_8 are alternatives and limited, thus cluster 7 in which C_{15} is located should be designed as a customized module	d
5	Values of S_4 have a large number of alternatives, thus cluster 1 should be designed as a personized module	e
6	Adaptable interfaces are proposed to connect components in different clusters	f

ences for product specification combinations using big sales data.

- The correlations of components originating from customers/market preferences with product specifications are identified by defining the relationship between specifications and components.
- The proposed method can accurately and comprehensively determine the correlation degree of product specifications and components from big sales data analysis, rather than defining the correlation degree using small amounts of data from experts and/or customers.
- The six rules proposed for supporting design decision marking can potentially facilitate axiomatic design, adaptable design, product family design, product platform design, and open architecture product design.

Although the effectiveness of the newly proposed method has been illustrated using a case study of EVs, in addition to specification correlations, other metrics such as the size, shape, and materials of components for product modularity evaluations should be considered for industrial applications. Additionally, the accuracy of the design recommendations provided by the proposed method is highly influenced by the following factors, which should be considered.

(1) Quality of collected data. For a given category of products with certain similarities, a comprehensive market survey should be conducted to obtain the number of product sales and combinations of product specifications.
(2) Calculation of correlation coefficients. The relationships among product specifications due to customer preferences could be linear and nonlinear. Both linear and nonlinear correlation coefficients should be considered in practice for the correlation analysis of the product specifications.
(3) Design objection-related scenario selection. The selection of the three types of design scenarios provided in Section 2.3 should be conducted based on specific design objectives for the relation transformation.
(4) The determination of thresholds t_1 and t_2 in Section 2.5. In the proposed work, two thresholds, t_1 and t_2, were defined empirically based on the designers' experience. Inappropriate value determination of t_1 and t_2 can lead to inappropriate recommendations for common, customized, and personalized modules.

Nevertheless, this work provides the first attempt at specification analysis using big sales data to support modular product design decisions. More research and application efforts are necessary to facilitate further product design using big sales data. Future research activities in this research scope include but are not limited to the followings:

- Investigation of evolution trends of product specifications in the market using historical data on product sales.
- Calculating functional correlations among product specifications to better facilitate decoupling strategy planning for a complex product or system.
- The effects of physical constraints of both the component features and manufacturer technique capability make the proposed method more specific for design decision support.

Authors' Contributions

JZ and PG were in charge of the whole trial; JZ, BL and QP wrote the manuscript; BL assisted with collecting and analyzing raw data from public website. All the authors read and approved the final manuscript.

Authors' Information

Jian Zhang, born in 1982, is currently a full professor at *Department of Mechanical Engineering, Shantou University, China*. He received his PhD degree from *University of Calgary, Canada*, in 2014. His research interests include design theory and methodology, big data in design and manufacturing, mass customization and personalization, and energy storage system design.
Bingbing Li, born in 1993, is a graduate student at *Shantou University, China*. She received her bachelor's degree on mechanical engineering from *Henan Polytechnic University, China*, in 2017.
Qingjin Peng, born in 1957, is currently a full professor at *Department of Mechanical Engineering, University of Manitoba, Canada*. He received his PhD degree from *University of Birmingham, UK,* in 1998. His research interests include intelligent manufacturing, product assembly and sustainability, design for personalization, modelling and simulation of manufacturing systems.
Peihua Gu, born in 1953, is currently a full professor at *School of Mechanical Engineering, Tianjin University, China*. He is a fellow of *Canadian Academy of Engineering* and a fellow of *International Academy of Production Engineering-CIRP*. He received his PhD degree from *McMaster University, Canda*, in 1990. His research interests include adaptable design, robust design, and computer-aided design/computer-aided manufacturing (CAD/CAM).

Funding

Supported by National Key R&D Program of China (Grant No. 2018YFB1701701), Sailing Talent Program, Guangdong Provincial Science and Technologies Program of China (Grant No. 2017B090922008) , and Special Grand Grant from Tianjin City Government of China.

Competing interests

The authors declare no competing financial interests.

Received: 2 April 2021 Revised: 30 December 2022 Accepted: 6 January 2023
Published online: 13 February 2023

References

[1] K T Ulrich, S D Eppinger. *Product design and development*. New York: McGraw-Hill Education, 2004.
[2] Q L Xu, S K Ong, A Y C Nee. Function-based design synthesis approach to design reuse. *Research in Engineering Design,* 2006, 17(1): 27-44.
[3] J C Aurich, C Fuchs, C Wagenknecht. Life cycle oriented design of technical product-service systems. *Journal of Cleaner Production*, 2006, 14(17): 1480-1494.
[4] M B Beverland, M T Ewing, M J Matanda. Driving-market or market-driven? A case study analysis of the new product development practices of Chinese business-to-business firms. *Industrial Marketing Management*, 2006, 35(3): 383-393.
[5] W Kuncoro, W O Suriani. Achieving sustainable competitive advantage through product innovation and market driving. *Asia Pacific Management Review*, 2018, 23(3): 186-192.
[6] P Gu, M Hashemian, A Y C Nee. Adaptable design. *CIRP Annals-Manufacturing Technology*, 2004, 53(2): 539-557.
[7] Y Koren, S J Hu, P H Gu, et al. Open-architecture products. *CIRP Annals-Manufacturing Technology*, 2013, 62(2): 719-729.
[8] J Zhang, G Xue, H L Du, et al. Enhancing interface adaptability of open architecture products. *Research in Engineering Design*, 2017, 28(4): 545-560.
[9] J Jiao, T W Simpson, Z Siddique. Product family design and platform-based product development: A state-of-the-art review. *Journal of International Manufacturing*, 2007, 18(1): 5–29.
[10] T J Marion, H J Thevenot, T W Simpson. A cost-based methodology for evaluating product platform commonality sourcing decisions with two examples. *International Journal of Production Research*, 2007, 45(22): 5285-5308.
[11] T W Simpson. Product platform design and customization: Status and promise. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 2004, 18: 3-20.
[12] J Zhang, D Xue, P Gu. Adaptable design of open architecture products with robust performance. *Journal of Engineering Design*, 2015, 26(1-3): 1-23.
[13] J Zhang, P Gu, Q Peng, et al. Open interface design for product personalization. *CIRP Annals-Manufacturing Technology*, 2017, 66(1): 173-176.
[14] H J Thevenot, T W Simpson. A comprehensive metric for evaluating component commonality in a product family. *Journal of Engineering Design*, 2007, 18(6): 577-598.
[15] J Bonvoisin, F Halstenberg, T Buchert, et al. A systematic literature review on modular product design. *Journal of Engineering Design*, 2016, 27(7): 488-514.
[16] J Zhang, B B Li, A Simeone, et al. Design decision based on dependencies among product specifications. *Proceedings of the 13th International Conference on Axiomatic Design*, Sydney, Australia, October 18-20, 2019: 30100012.
[17] Z Y Liu, S G Zhou, J R Tan, et al. Multi parameter correlation analysis and prediction method of product performance based on multi criteria correction. *Journal of Mechanical Engineering*, 2013, 49(15): 105-114.
[18] J Sangjin, A Oyku, T W Simpson. A method to evaluate direct and indirect design correlations between components in a product architecture. *Research in Engineering Design*, 2018, 29(3): 507-530.
[19] J Zhang, A Simeone, P Gu, et al. Product features characterization and customers' preferences prediction based on purchasing data. *CIRP Annals–Manufacturing Technology*, 2018, 67(1): 149-152.
[20] O Kramer. On missing data hybridizations for dimensionality reduction. In: *Hybrid metaheuristics*. Berlin Heidelberg: Springer, 2013.
[21] I H Witten, E Frank, M A Hall, et al. *Data mining: Practical machine learning tools and techniques*. Massachusetts: Morgan Kaufmann, 2016.
[22] R A S Fisher, J H Bennett. *Statistical methods, experimental design and scientific inference*. Cambridge: Oxford University Press, 1990.
[23] L Myers, M J Sirois. Spearman correlation coefficients, differences between. In: *Encyclopedia of statistical sciences*. New York: John Wiley & Sons Inc., 2006.
[24] H B Barlow. Unsupervised learning. *Neural Computation*, 1989, 1(3): 295-311.
[25] K A Ciesielski, T Mandat. Unsupervised machine learning in classification of neurobiological data. In: *Intelligent methods and big data in industrial applications*. Cham: Springer, 2019.
[26] G Erixon. Modular function deployment—a method for product modularization. *The Royal Institute of Technology*: *Stockholm*, 1998.
[27] F Börjesson, K Höltä-Otto. A module generation algorithm for product architecture based on component interactions and strategic drivers. *Research in Engineering Design*, 2014, 25(1): 31-51.
[28] K Wegener, S Andrew, A Raatz, et al. Disassembly of electric vehicle batteries using the example of the audi Q5 hybrid system. *Procedia CIRP*, 2014, 23: 155-160.