

Research paper

A new data-driven design methodology for mechanical systems with high dimensional design variables

Xianping Du, Feng Zhu*

Department of Mechanical Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA



ARTICLE INFO

Keywords:

High-dimensional design variables
 Vehicle crashworthiness
 Structural design
 Data mining
 Critical parameters identification (CPI)
 Design domain reduction (DDR)

ABSTRACT

Complicated engineering products such as cars with a large number of components can be regarded as big data systems, where the vast amount of dependent and independent design variables must be considered systematically during the product development. To design such a system with high-dimensional design variables, this study aims at developing a novel methodology based on data mining theory, and it is implemented through designing a crashworthy passenger car, which is a multi-level (system – components) complicated system. Decision tree technique was used to mine the crash simulation datasets to identify the key design variables with most significant effect on the vehicular energy absorption response and determine the range of their values. In this way, the design space can be significantly reduced and the high-dimensional design problem is greatly simplified. The results suggest that the data mining based approach can be used to design a complicated structure with multiple parameters effectively and efficiently. Compared with the traditional design method, the new approach could simplify and speed up the design process without significant influence on the accuracy.

1. Introduction

Design of complicated mechanical products requires complex system engineering, where a large number of dependent and independent design variables must be considered systematically [5,30]. These design variables in such high dimensional systems may have very complex interrelationship and be of multi-level and multi-physical nature. For example, the safety (crashworthiness) design of a vehicle can be considered as a multi-level design problem, in which crash energy management is implemented by thousands of coupled and uncoupled geometric and material design variables at system, component and sub-component levels. Many design factors, such as load path, structural deformations and component collapse sequence, as well as vehicle size, weight, geometry, etc. must be considered at the same time. To design such a complicated system, efficient methods are necessary to identify the relationship among these variables and derive the right design workflow from vast amount of data. Another example is design of battery, which is a multi-physical system. The overall performance of a battery is governed by a large number of mechanical, thermal, chemical and electrical design variables. These variables may have strong coupling effect and influence one another. To achieve an effective and efficient design, it is imperative to identify such complex coupling effect, and this could be very challenging when considering a high dimensional system.

To resolve the aforementioned issues associated with high dimensional design space or dataset, two strategies, namely critical parameters identification (CPI) and design domain reduction (DDR) can be applied. CPI is to identify those design variables which have strong effect on the system performance. Based on these key variables, DDR is then performed to reduce the range of values of the variables to further decrease the design space [20,31]. Some researches were conducted on CPI through sensitivity analysis [7,22,26], analysis of variance (ANOVA) [9], and other methods [10,13,14,25]. The examples for DDR includes the recursive feasible-infeasible segment bisection algorithm proposed by Lanzi et al [17], the probability of feasibility method by Basudhar et al [2], and other approaches [3,11,23]. However, most of these methods for CPI or DDR have no capability to reveal the coupling effect of the design variables from big datasets.

As one of the most promising approaches to overcome this limitation, data mining has been increasingly applied to engineering design to improve the design process for complicated systems. Zheng et al screened the key design parameters for vehicle weight minimization using fuzzy rough set method [31]. Liu et al reduced initial design space for an optimization problem using classification and regression tree (CART) [18]. Shi et al. used a similar technique to reduce the design domain for a multi-objective optimization and demonstrated the performance of this method using a speed reducer and vehicle design [20]. Chen et al. applied the classification, association and clustering method

* Corresponding author at: #Mailing address: 600 S Clyde Morris Blvd, Daytona Beach, FL 32114, USA.
 E-mail addresses: DUX1@my.erau.edu (X. Du), zhuf1@erau.edu (F. Zhu).

to achieve the similar goals [6]. These studies considered either CPI or DDR in high dimensional design problems but did not implement them simultaneously.

To resolve this issue, in this study, a decision tree – based data mining approach is proposed and applied to the engineering product design. Decision tree is a tree-like graph to identify the key variables in a large design space or dataset, and classify their effect on the system performance. Meanwhile, according to the design objectives, the range of the values of these key parameters can be determined quantitatively. In this way, the further design can focus on these “high impact” variables with smaller range of values, and thus CPI and DDR are achieved at the same time. The new approach is implemented and applied to the crashworthiness design of a passenger car as a case study. Its performance is verified through comparison with a conventional response surface – based design method.

2. Theoretical frame of the new design methodology

2.1. Theoretical analysis on data mining for CPI and DDR

A data mining technique, namely Decision Tree Method [12], was used to implement CPI and DDR in this study. A decision tree is a tree- or flowchart-like structure in which each internal node represents a “test” on an attribute (i.e. a design variable in the current case), each branch represents the outcome of the test, and each leaf node represents a class label for the performance of that design. The paths from root to leaf represent classification or design rules. In the design of a complicated system, a large number of design alternatives are generated by accounting the combinations of all design variables and their possible values. These design alternatives are computed with Finite Element (FE) analysis subjected to the pre-defined constraints. Without CPI and DDR, the simulation dataset would be of huge size and the computational cost is not affordable. Using a decision tree, the attributes or design variables associated with the selected decision making/classification path can be deemed as the key design variables, and thus CPI is achieved by performing the design with these variables only. Likewise, the nodes of the decision tree specify the range of design variable's values to implement DDR. An example will be presented in Section 4 to illustrate the design procedure in detail.

There are several algorithms available to create a decision tree [12]. Here an algorithm, termed as J48 [4,19], was used. It is an extension of the basic decision tree generation algorithm C4.5, which uses the gain ratio of entropy as the node splitting criterion. The gain ratio describes the degree of purity in the resulting partitions, and eliminates the bias of selecting partition variables with a large amount of levels. If a design variable is tested on a non-leaf node, the gain ratio with respect to this parameter represents the degree of subset purity if classified by this variable. Hence, at each non-leaf node, the design variable with maximum gain ratio is selected to partition the dataset. Splitting value of this variable is determined by the resulting subsets to ensure that each resulting subsets are as “pure” as possible by testing each variable level. A decision tree with a large number of variables is constructed by recursions until any one of the following terminating conditions is true: (1) all tuples of the whole dataset is in the same class; 2) the new subset is empty; 3) all design variables are tested.

In addition to J48, a pruning operation is performed to avoid the redundant nodes of decision tree [12]. After the decision tree is constructed and pruned, CPI and DDR can be implemented automatically by following the top-down design path.

2.2. Outline of the data mining procedure

The basic procedure of the data mining-based design method is shown in a flowchart in Fig. 1. In the first step, the design problem is analyzed and the design objective is determined. Likewise, the design variables are further identified and the range of their values are

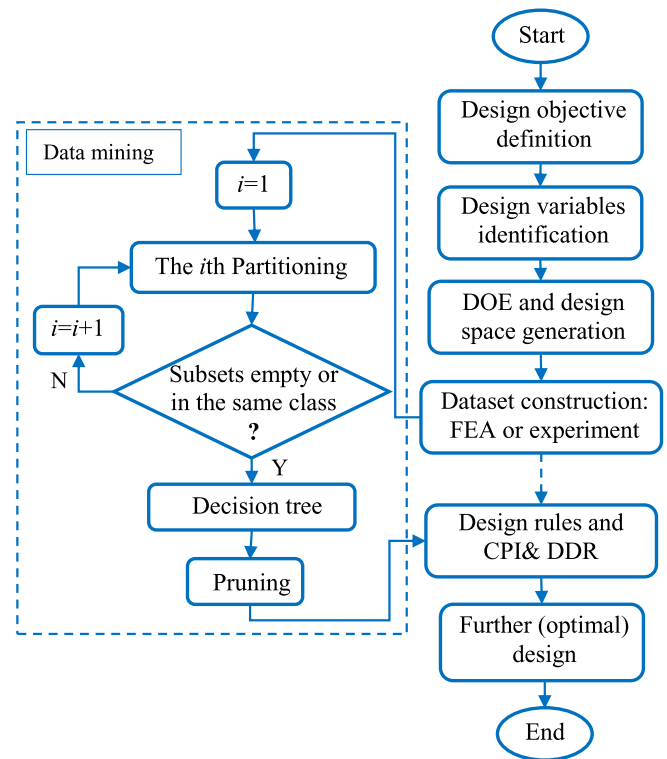


Fig. 1. Flowchart of the data mining method to design a system with high-dimensional design variables.

determined. Then, in the design space, a large number of design of experiments (DOEs) are built. Each design is computed through FEA and the simulation results form a large design dataset. After that, data mining is performed on the simulation dataset to construct a decision tree, with which CPI and DDR can be implemented simultaneously to reduce the design space. Likewise, useful inter-relationships among design variables and decision making rules can be derived. Based on this information and reduced design space, the further optimal design can be achieved faster and at lower cost.

In the data mining process, the recursive partitioning on the simulation dataset (the initial design space) is conducted to construct the decision tree. The gain ratio of entropy is first calculated for the whole dataset to determine the first ($i = 1$) partitioning. If all subsets after this operation are empty or belong to the same class, the decision tree construction is complete. Otherwise, a new round of partitioning is performed ($i = i + 1$) on the impure subsets and each of these impure subsets includes more than one class. The procedure is repeated in this recursive way until all subsets are empty or in the same class. After the decision tree is built, it is pruned to obtain more concise results. Using this decision tree, design rules are derived, and then CPI and DDR are carried out to achieve the final design.

3. Case study: Crashworthiness design of vehicle structure

In this section, the safety (crashworthiness) design of a vehicle is used as a case study to demonstrate the feasibility of this new approach. The results of the current approach will be compared with the design results based on a commonly used response surface - based method to evaluate its performance.

3.1. The problem description

Vehicle crashworthiness connotes a measure of plastic deformation of vehicular structure and its maintenance of a sufficient survival space for its occupants during crashes resulting in reasonable deceleration

Download English Version:

<https://daneshyari.com/en/article/6961365>

Download Persian Version:

<https://daneshyari.com/article/6961365>

[Daneshyari.com](https://daneshyari.com)