Journal of Cleaner Production 68 (2014) 189–199

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

# Demand Trend Mining for Predictive Life Cycle Design

Jungmok Ma [a], Minjung Kwak [b], Harrison M. Kim [a,*]

[a] Enterprise Systems Optimization Laboratory, Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, 104 S. Mathews Ave., Urbana, IL 61801, USA

[b] Department of Industrial and Information Systems Engineering, Soongsil University, Dongjak-Gu, Seoul, Republic of Korea

ARTICLE INFO

*Article history:*
Received 30 April 2013
Received in revised form
11 November 2013
Accepted 5 January 2014
Available online 22 January 2014

*Keywords:*
Demand Trend Mining
Product and design analytics
Data-driven product design
Product life cycle design
Decision tree

ABSTRACT

The promise of product and design analytics has been widespread and more engineering designers are attempting to extract valuable knowledge from large-scale data. This paper proposes a new demand modeling technique, Demand Trend Mining (DTM), for Predictive Life Cycle Design. The first contribution of this work is the development of the DTM algorithm for predictability. In order to capture hidden and upcoming trends of product demand, the algorithm combines three different models: decision tree for large-scale data, discrete choice analysis for demand modeling, and automatic time series forecasting for trend analysis. The DTM dynamically reveals design attribute pattern that affects demands. The second contribution is the new design framework, Predictive Life Cycle Design (PLCD), which connects the DTM and data-driven product design. This new optimization-based model enables a company to optimize its product design by considering the pre-life (manufacturing) and end-of-life (remanufacturing) stages of a product simultaneously. The DTM model interacts with the optimization-based model to maximize the total profit of a product. For illustration, the developed model is applied to an example of smart-phone design, assuming that used phones are taken back for remanufacturing after one year. The result shows that the PLCD framework with the DTM algorithm identifies a more profitable product design over a product life cycle when compared to traditional design approaches that focuses on the pre-life stage only.
© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

### 1.1. Demand Trend Mining in design analytics

Product design analytics or data-driven product design is emerging as a promising area by bridging benefits of large-scale data and product design decisions. With the popularity of social network and web devices, a large volume of data which has a characteristic of complexity, timeliness, heterogeneity, and lack of structure (Labrinidis and Jagadish, 2012) are being generated every day. Although the necessity of large-scale data analysis for product design is now being recognized broadly, only a few researchers have attempted to analyze large-scale data in the context of product and design analytics (Tucker and Kim, 2008, 2011b; Van Horn et al., 2012). This paper proposes Demand Trend Mining (DTM) as one of the analysis tools for large-scale data in order to capture the trend of demand as a function of design attributes. The DTM is a dynamic and adaptive model in that it mines the underlying changes of concept drift from time series data and builds a predictive model based on the changes. The model shows that it can realize Predictive Life Cycle Design which encompasses both the *pre-life* (i.e., manufacturing) and *end-of-life* (i.e., remanufacturing[1] and recycling) stages.

### 1.2. Remanufacturing and life cycle design

Remanufacturing has been a new profit opportunity for original equipment manufacturers (OEMs). Caterpillar, Xerox, and Sony are among the OEMs who have successfully taken this new opportunity (Hucal, 2008; King et al., 2006; Parker and Butler, 2007). In remanufacturing, used products are restored to a like-new condition and are given another life in the market. Remanufacturing can bring larger profits over the span of a product from an initial investment at low additional costs, typically 40%–65% less than new product costs because it reutilizes the materials and the value added to a product in its initial manufacturing (Pearce, 2008; Lund, 1984).

Remanufacturing also enables OEMs to improve their environmental performance. As awareness of environmental issues

* Corresponding author. Tel.: +1 217 265 9437; fax: +1 217 244 5705.
*E-mail addresses:* hmkim@illinois.edu, hmkim@uiuc.edu (H.M. Kim).

[1] In this paper, remanufacturing is used as an umbrella term which encompasses reuse, reconditioning, refurbishment, and cannibalization.

0959-6526/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jclepro.2014.01.026

increases, pressure from the public and policymakers have prompted OEMs to be responsible for the environmental impacts of their products. OEMs now need to extend their environmental efforts to encompass the entire life cycle of a product, from cradle (raw material extraction) to grave (end-of-life disposal). By remanufacturing a product, OEMs can reduce waste and minimize the need for raw material to make new products. It is known that remanufactured products (hereinafter *reman product*) can save up to 90% of the environmental impact of entirely new products (Charter and Gray, 2007; Parker and Butler, 2007).

In order for successful remanufacturing, design for life cycle (or life cycle design) is key for OEM remanufacturers. Product design determines not only the current profit from the pre-life stage but also the future profit from the end-of-life stage (Newcomb et al., 1998; Kwak and Kim, 2010; Zhao and Thurston, 2010). Therefore, to maximize the total profit from the entire life cycle of a product, OEM remanufacturers must optimize their design decisions considering both stages together.

### 1.3. Challenges and contributions

The main challenge in life cycle design is that there is a significant time gap (i.e., usage-life) between the pre-life and end-of-life stages. As illustrated in Fig. 1, suppose that the decision maker is at time $t^{prelife}$ (design stage), and the selling point of new product is $t^{first}$. In this research, it is assumed that the time gaps between $t^{prelife}$ and $t^{first}$, and $t^{eol}$ and $t^{second}$ are known. Also, it is assumed that the usage-life is $h$, remanufacturing will occur at time $t^{eol}$, and the remanufactured products will be sold at the market at time $t^{second}$. For instance, the typical usage-life of cell phones and laptops is known as 1.5 years (Cellular-Recycler, 2011) and 4 years (Deng et al., 2009), respectively. Considering rapid changes in technology and customer preferences, such a time gap between pre-life and end-of-life stages implies that life cycle design should consider and satisfy two sets of customer needs at the same time, i.e., needs for new products at the present and needs for reman products in the future. Although many demand models have been presented for capturing current demands at the new-product market (hereinafter *new market*), very few models are available for forecasting future demands at the remanufactured-product market (hereinafter *reman market*). Moreover, little research has been presented that combines a dynamic demand model with life cycle design, which considers the time gap and transforms a trend of customer preferences to projected demands.

Another challenge is uncertainty of returned products in terms of quantity, timing, and condition. Fig. 1 shows material flow starting from material extraction to part manufacturing, product assembly, recovery and disposal. The scope of the problem is clearly defined using solid arrows. In this paper, recovery options are categorized as material, part, and product levels. Product level recovery (e.g., reuse and reconditioning) only requires some minor value-added operations including polishing, cleaning, and lubricating. Part level recovery (e.g., cannibalization and refurbishment) needs disassembly as well as parts conditioning and change. Material level recovery (e.g., recycling) is usually conducted by recyclers and raw materials are recovered by shredding and refining. There are three possible cases that require corporations' end-of-life decisions: initial returns, returns within warranty period, and take-back program. The initial returns are caused by changes of purchase decisions in a short period of time. The returns within warranty period are induced by defects in any time. The focused case, take-back program, aims at boosting sales with re-purchasing contracts of sold products within specified period. In this case, the amount and condition of returned products should be considered in a model.

We propose the DTM algorithm which is depicted in Fig. 2 to systematically tackle some challenges: extracting valuable knowledge from large-scale data, building a demand model from the mined knowledge, and predicting a target demand in the future. The requirements for overcoming the challenges include 1) utilization of large-scale data, 2) estimation of demands, and 3) realization of demand trends over time. In order to fulfill these requirements, the DTM algorithm utilizes and combines three different models: discrete choice analysis (DCA), decision tree, and automatic time series forecasting. If $t = 1$ to $t = n$ data are available and $t = h$ ahead demand is needed, then the DTM provides a way to estimate demand at $t = n + h$ as shown in Fig. 2. To combine the DCA and decision tree, a class variable of a decision tree model is proposed to be expressed as utility. Also the concept of generational difference is adopted for a prevention of missing values and smooth forecasting in product design.

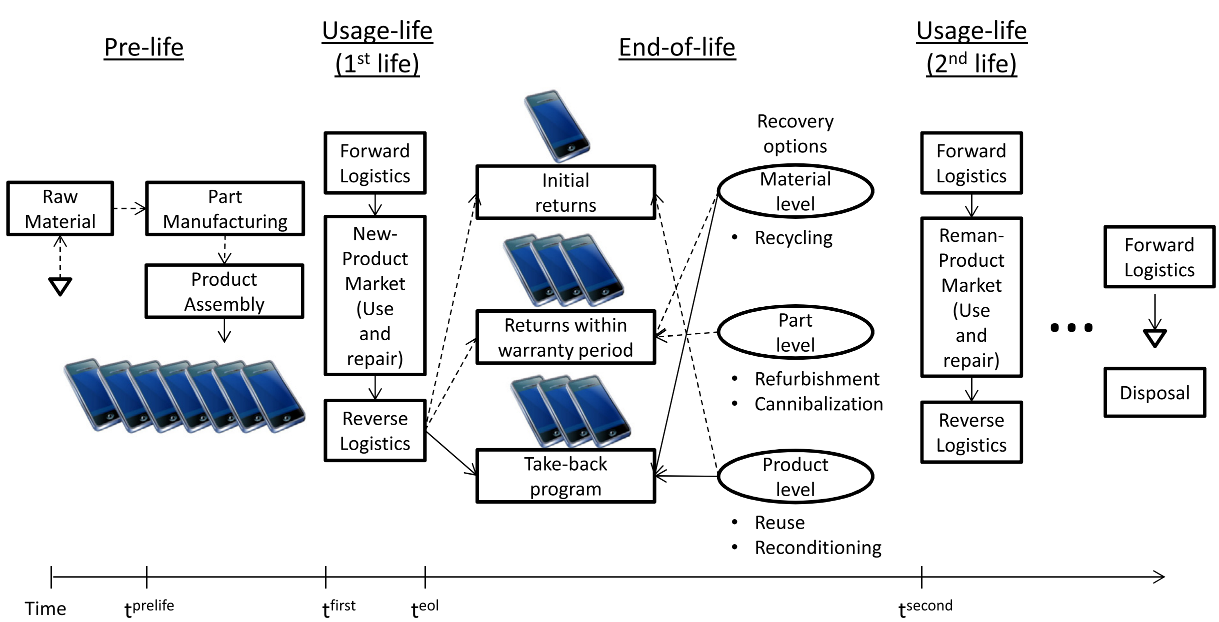


**Fig. 1.** Closing the loop of product life cycle and scope of the problem (solid arrow).

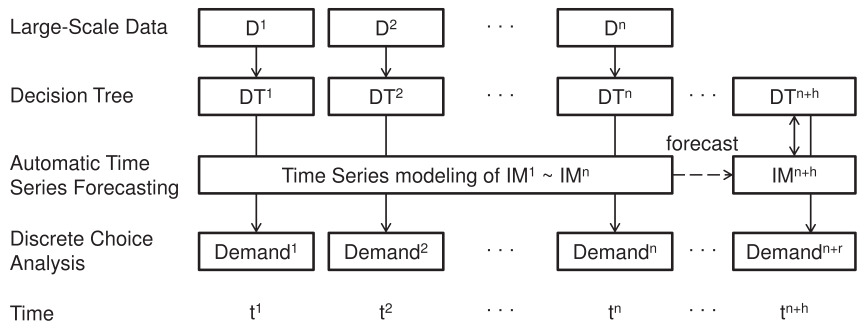


**Fig. 2.** Demand trend mining algorithm.

Using the DTM algorithm, Predictive Life Cycle Design (PLCD) can be finally implemented. The proposed PLCD framework enables engineering designers to optimize target product design by considering the pre-life and end-of-life stages of a product simultaneously. The identified product design will maximize the total profit over the entire product life cycle. Fig. 3 provides an overview of the PLCD framework. The dotted box represents the DTM model. The remaining components represent the optimal life cycle design or optimization-based model. The framework optimizes the product attributes as well as the selling prices and production quantities of new and reman products. For illustration, the developed model is applied to an example of smart-phone design, assuming that the available products from initial sales of the pre-life will return for remanufacturing after one year of usage, according to a take-back contract.

The rest of the paper is organized as follows. Section 2 reviews relevant literature, and Section 3 describes the detailed steps of overall methodology. Section 4 presents an illustrative case study of smart-phone design. Section 5 concludes the paper with suggestions for future research.

## 2. Background and previous literature

### 2.1. Demand Trend Mining

Modeling demand and customer preferences are critical for assessing the profit of a product. Under the framework of Predictive Life Cycle Design (PLCD), the decision maker should consider two markets at the design stage, i.e., the current market for new products and the future market for reman products. Considering the time gap between pre-life and end-of-life stages, customer preferences in the two markets are likely to be different. The DTM thus aims to construct two demand models: one for new products and the other for reman products.

Two widely used demand analysis techniques in product design are discrete choice analysis (DCA) (Wassenaar and Chen, 2003; Wassenaar et al., 2005) and conjoint analysis (Moore et al., 1999; Grissom et al., 2006). While both techniques can capture customers choice behavior and model related demands, they resort to direct customer interactions (e.g., survey) and have a limited capability to use large-scale data due to the statistical assumptions (Tucker and Kim, 2009).

Decision tree in data mining is an alternative model for customer preferences in product design. Since the decision tree algorithm can deal with large-scale massive data, it was proposed to generate product concepts for engineering designers (Hall et al., 1998; Tucker and Kim, 2009). However, very little research was conducted on demand analysis with the decision tree in the field of product design and other system design (Tucker and Kim, 2011b; Yu et al., 2010).

In order to capture trends of demand, dynamic demand models should be constructed instead of static demand models. Dynamic models do not assume that demand models that were once built would remain the same over time. Böttcher et al. (2008) suggested

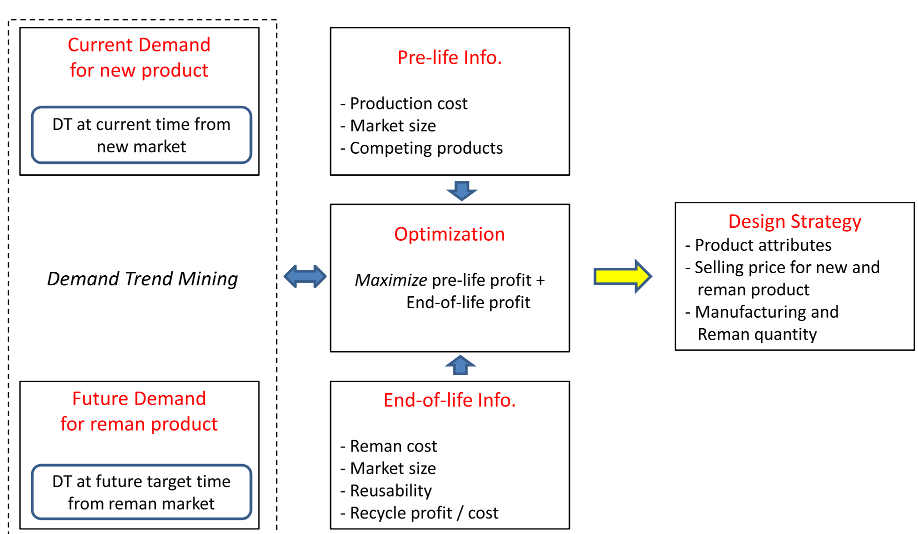


**Fig. 3.** Summary OF PLCD framework.

that decision tree can be built based on predicted values of interestingness measure (IM). IM is a term for “various measures devised for evaluating and ranking discovered patterns produced by the data mining process” (McGarry, 2005). To trace the trend of IMs, a polynomial regression model was utilized. Tucker and Kim (2011b) suggested the adoption of the time series analysis technique, Holt–Winters exponential smoothing model, which is a more complex modeling technique with time-variant data. However, there exist different classes of exponential smoothing, which means the Holt–Winters model is just one of its family and engineering designers should choose right one among them. At the same time, designers are required to determine many different parameters and initial states for the Holt–Winters model. The DTM algorithm adopted the Hyndman's automatic time series forecasting algorithm (Hyndman et al., 2008; Hyndman and Khandakar, 2008). This algorithm includes the automatic optimization process for model selection, parameter setting, and initial state estimation with the *innovations* formulation of state space models.

The proposed DTM combines the merits of aforementioned three different models: DCA for demand modeling, decision tree for large-scale data, and automatic time series forecasting for trend analysis. The decision tree algorithm, C4.5, models customer preferences from large-scale data, and by formulating a class variable as utility, the resulting decision tree models can estimate market shares from the DCA, specifically logit choice probability in multinomial logit (MNL) model. Automatic time series forecasting provides predicted IMs, and trend reflected demand is estimated from the target time decision tree.

Table 1 provides a summary of the MNL and C4.5 (Tucker and Kim, 2011b; Tucker et al., 2009). The MNL model starts from a random utility model where the true utility consists of the observable utility and the unobservable random part. In the MNL, the random part is assumed as independent and identically distributed extreme value, and the choice probability is given by the logit choice probability. The C4.5 is based on the information theory. Entropy, a measure of disorder or complexity, is calculated, and the decision tree is built in the direction of minimizing the entropy.

### 2.2. Predictive Life Cycle Design

Design for life cycle or life cycle design focuses on the fact that decisions made at the design stage affect all phases of a product's life cycle (i.e., material extraction, manufacturing, usage, and end-of-life recycling and disposal). In many previous studies, it has been emphasized that the design stage determines 70–85% of a product's total life cycle cost and environmental impact (Fixson, 2004; Duverlie and Castelain, 1999; Seo et al., 2002). Therefore, life cycle design is aimed at proactively dealing with economic and environmental issues during the early design stage when the potential for affecting results is the greatest. Since little research has been conducted for the economic perspective in comparison with the matured environmental perspective over entire life cycle (Hundal, 2001; Kwak, 2012), only the economic side of product design is considered in this paper. However, the economic benefits from the end-of-life stage that are traditionally hidden source of profit can lead to environmentally friendly design by considering end-of-life processes.

Some researchers (Lye et al., 2001; O'Shea, 2002; Holt and Barnes, 2010) have developed a holistic design approach that considers various concerns of all life cycle stages in an integrated manner. However, a more popular approach has been to develop design principles for improving a specific life cycle stage. Design for recovery, design for remanufacturing, design for disassembly, and design for recycling are among the principles of life cycle design. Focusing on the end-of-life stage, they seek to identify optimal product design to reduce the cost of recovery and/or increase the profit associated with recovery.

Rose et al. (2000, 2002) suggested a classification scheme for helping designers predict appropriate recovery strategies for a product, so that the designers can redesign products to move toward a higher level of reuse. Mangun and Thurston (2002) developed a model for designing a product portfolio that incorporated part reuse through refurbishment. Given multiple market segments with varying requirements for environmental impact, production cost, and reliability, they attempted to determine the optimal product design for each segment in order to maximize the total utility of the portfolio. Kwak and Kim (2010, 2011) introduced a framework for analyzing how product design affects end-of-life recovery and what architectural characteristics are desirable for higher recovery profit.

One limitation of these previous methods, however, is that the design implications on the pre-life and end-of-life stages have been considered separately. Product design has been optimized for each of the stages, but not for the stages together due to the lack of available demand forecasting models. By the proposed DTM algorithm, both stages now can be considered together. This is why the framework is called *Predictive* Life Cycle Design. Two exceptions can be found in Zhao and Thurston (2010) and Kwak and Kim (2013a). Both developed a mathematical model to determine the optimal product design that maximizes the profits from both initial sales and end-of-life recovery. They showed that the total profit can be maximized when both ends of the product life cycle are considered at the same time. However, the prediction and reflection of demand trends in the market were not incorporated.

## 3. Methodology

This section describes detailed steps for the DTM and PLCD. Fig. 4 shows the overall framework of the PLCD which has two components: DTM and optimal life cycle design. Although the general description of the DTM algorithm is described in Figs. 2 and

**Table 1**
Overview Of Mnl and C4.5 (Tucker and Kim, 2011b; Tucker et al., 2009).

| | MNL | C4.5 |
|---|---|---|
| Assumption | - Random Utility Model<br>$U_{nj} = V_{nj} + \epsilon_{nj}$<br>$\epsilon_{nj} \sim$ *iid extreme value*<br>$j$: choice alternative<br>$n$: decision maker | - Information Theory (Information Entropy)<br>$Entropy(D) = -\sum_{i=1}^{k} P_i \cdot \log_2 P_i$<br>$D$: *data set*<br>$k$: *number of class variables within the data set*<br>$P_i$: *probability of class variable i* |
| Choice Probability & Split Criterion | - Logit Choice Probability<br>$P_{ni} = \frac{\exp(V_{ni})}{\sum_j \exp(V_{nj})}$<br>Decision maker $n$ choose alternative i over alternative j (i $\neq$ j) | - Gain Ratio<br>$Gain\ Ratio(X) = \frac{Entropy(D) - \sum_{j=1}^{n} \frac{\lvert D_j \rvert}{\lvert D \rvert} \cdot Entropy(D_j)}{-\sum_{j=1}^{n} \frac{\lvert D_j \rvert}{\lvert D \rvert} \cdot \log_2 \frac{\lvert D_j \rvert}{\lvert D \rvert}}$<br>$X$ : *attribute*<br>$n$ : *number of outcomes for a given attribute* |

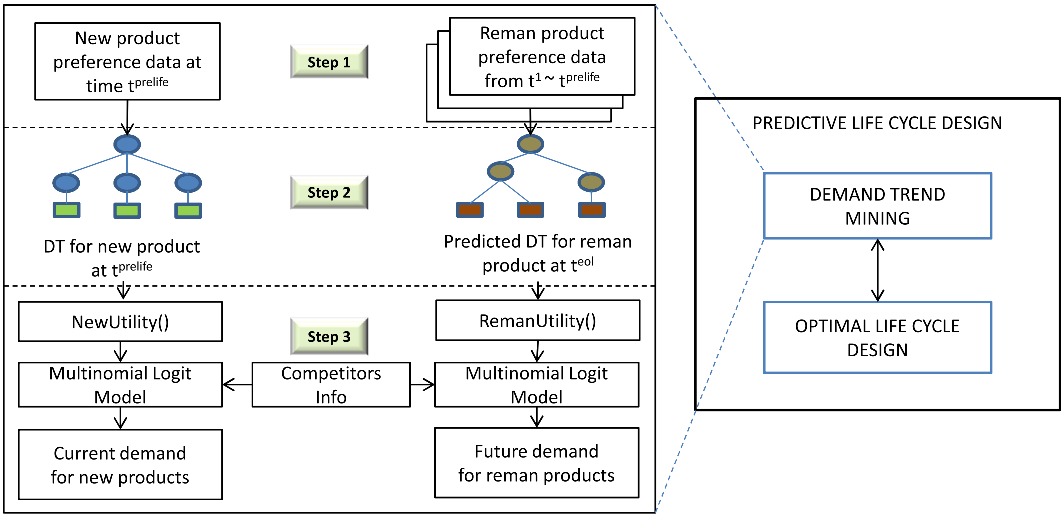


**Fig. 4.** Framework Of PLCD.

4 provides more detailed steps of the DTM, especially in the framework of the PLCD.

### 3.1. Modeling of demand trend

As illustrated in Fig. 1, suppose that the decision maker is at time $t^{prelife}$, and the selling point of new product is $t^{first}$. It is assumed that the time gaps between $t^{prelife}$ and $t^{first}$, and $t^{eol}$ and $t^{second}$ are known. Also, it is assumed that the usage-life is $h$, remanufacturing will occur at time $t^{eol}$, and the remanufactured products will be sold at the market at time $t^{second}$. Thus, the PLCD framework starts from the DTM which estimates the market demands at time $t^{first}$ and time $t^{second}$ for new and reman products, respectively. The DTM algorithm in Fig. 2 is divided as 3 Steps in the following subsections in detailed description. Step 2 covers decision trees and automatic time series prediction, which are components of Preference Trend Mining. The demand modeling with discrete choice analysis is depicted in Step 3.

#### 3.1.1. Step 1: data collection

In the first step, two data sets are collected to capture trends of demand in the market. First, the customer preference data for new products are collected at the current time $t^{prelife}$ in Fig. 1. This data is used for capturing market demand at the pre-life design stage. Second, the historical and the current preference data for reman products are collected to predict market demand at the end-of-life stage. The preference data from time $t^1$ to $t^{prelife}$ in the reman market are used to mine underlying demand trends and estimate the market demand at time $t^{second}$.

Table 2 shows the basic data structure with an example of smart-phone design. The data comprises of two parts: a set of attributes and a class variable. Attributes are product features, and class variables are outputs or responses that we are interested in. In this research, the degree of customer preference or the customer utility on a discrete scale is used as the class variable. It can be either stated data from a survey or revealed data from online reviews. By having utility as the class variable, demand modeling is allowed in Step 3. Each attribute has its own levels; for example, the attribute camera pixel has two different levels, e.g., 8 or 16-MP.

In the case of attributes with significant improvement in their values, it is represented in a relative scale using the concept of generational difference (Kwak and Kim, 2013b). The generational difference can be acquired by comparing the generational gap between the target part and the latest cutting-edge part which corresponds to the maximum generation. For example, if 16-MP is the latest generation, then the generational difference is 0. If 8-MP is the previous generation, the generational difference is 1. The advantages of the generational difference include the prevention of missing values over time and the allowance of forecasting without specific levels so that emerging trends can be captured with various levels. The original Preference Trend Mining proposed by Tucker and Kim (2011b), which will be discussed in the next step, was not intended to deal with various levels, as the algorithm used fixed levels over time.

#### 3.1.2. Step 2: Preference Trend Mining

In the second step, the data sets collected in Step 1 are analyzed in order to reveal the link between product attributes and customer utility. For new and reman products, different analyses are conducted. The data for new products is analyzed using Quinlan's C4.5 decision tree algorithm (Quinlan, 1993) which is a static model. The generated decision tree can be transformed into a set of decision-tree-based rules, i.e., NewUtility(). Each path of the decision tree expresses a decision rule; given an attribute combination, the decision-tree-based rule provides an estimate of utility.

**Table 2**
Data structure (with example of smart-phone design).

| Smart-Phone Attribute | | | | | | | | | | Class |
|---|---|---|---|---|---|---|---|---|---|---|
| New product price | | Reman product price | | Screen size | | Memory | | Camera pixel | | Utility |
| \$199 | $Y_{11}$ | \$99.5 | $Y_{21}$ | 2.8″ | $X_{11}$ | 2 (16 GB) | $X_{21}$ | 1 (8 MP) | $X_{31}$ | 1 |
| \$299 | $Y_{12}$ | \$149.5 | $Y_{22}$ | 3.5″ | $X_{12}$ | 1 (32 GB) | $X_{22}$ | 0 (16 MP) | $X_{32}$ | 2 |
| \$399 | $Y_{13}$ | \$199.5 | $Y_{23}$ | 5.3″ | $X_{13}$ | 0 (64 GB) | $X_{23}$ | | | 3 |
| | | | | | | | | | | 4 |

The time series data for reman products is analyzed using the revised Preference Trend Mining (PTM) algorithm adopted by Tucker and Kim (2011b). The algorithm generates a predicted decision tree for the future time $t^{eol}$, which can provide a set of decision rules, i.e., RemanUtility(). Algorithm 1 shows the pseudo code for the PTM. $S_T$ is the time series data (from time $t^1$ to $t^{prelife}$) for reman products and $X$ is the set of attributes. The PTM algorithm is similar to the C4.5 algorithm in that it builds the decision tree based on the interestingness measure (IM). In both algorithms, the attribute with the maximum IM becomes the node for branching. The difference is in how to calculate the IMs.

Unlike the C4.5 using one aggregated data set, the PTM forecasts the IMs of the future time from the historical time series data. In Algorithm 1, the PTM starts from finding the IMs of all attributes $X$ from all previous data (line 3). Then, there are processes to predict the IMs at $t^{eol}$ using the IMs from $t^1$ to $t^{prelife}$ and assign the attribute with the maximum IM as the root node of the tree (line 5). The levels of the attribute then become branches. For each branch, the same processes are repeated for remaining attributes; i.e. the PTM checks which attribute has the maximum IM at $t^{eol}$ and iteratively splits a decision tree until it reaches termination criteria. After identifying all the leaf nodes, the algorithm returns the predicted decision tree.

**Algorithm 1.** Preference Trend Mining revised from Tucker and Kim (2011b)

1: **procedure** PTM ($S_T$)
2: **while** Termination criteria are met **do**
3: Find IM(X) for $S_T$ and Forecast IM(X) at $t^{eol}$
4: *If* $IM(X_i) = MAX\ IM(X)$ at $t^{eol}$
5: Then $X_i$ = root node, $X_i$ levels = branches
6: Find $IM(X)$ for $S_T$ and Forecast $IM(X)$ at $t^{eol}$ given selected branches
7: If $IM(X_i') = MAX\ IM(X)$ at $t^{eol}$
8: Then $X_i'$ = childnode, $X_i'$ levels = branches
9: Repeat 6, 7, 8
10: **end while**
11: Result class variable = leaf node
12: **return** Predicted decision tree
13: **end procedure**

To apply the PTM algorithm, three issues should be clarified. First, the decision maker should decide the IM to use. The IMs that are well known and widely used include Shannon's entropy, gini index, information gain, and gain ratio. Depending on the data and its characteristics, each measure has its own pros and cons (Harris, 2002). In this paper, the gain ratio was selected following the C4.5 algorithm although the approach can be generalized with the other IMs.

The second issue is about the forecasting engine for the IM prediction. Hyndman's exponential smoothing (Hyndman et al., 2008) and the Box–Jenkins model (Box and Jenkins, 1976; Naylor et al., 1972) are among the most popular and widely-used methods for time series forecasting. In the Hyndman's exponential smoothing, the time series data can be decomposed into four components, i.e., trend, seasonal, cycle, and irregular error. A total of 30 mathematical models are available, and the best model can be obtained using automatic time series forecasting algorithm (Hyndman et al., 2008; Hyndman and Khandakar, 2008). The Box-Jenkins model is another popular option. It applies an autoregressive moving average (ARMA) or an autoregressive integrated moving average (ARIMA) to fit the time series data. Exponential smoothing has value in that it is relatively simple and easy to understand though there is no general consensus about which one has a better prediction accuracy (Gooijer and Hyndman, 2006; Geurts and Ibrahim, 1975). In this research, Hyndman's exponential smoothing model, specifically the automatic time series forecasting method, is chosen as the forecasting engine.

The difference between the Holt-Winters model in the original PTM (Tucker and Kim, 2011b) and the automatic time series forecasting (Hyndman et al., 2008, 2002) in the DTM is that the former is just one of exponential smoothing family and requires many of user inputs, but the latter allows automated model selection, and parameters and initial state estimation among 30 different linear and nonlinear models for designers.

Termination criteria in decision-tree generation is another important issue. If all class variables are distributed homogeneously and no valid split is found, the process can be stopped. If the leaf node is reached and the class variables are not distributed homogeneously, the path can be removed or the dominant class variable over time can be selected.

#### 3.1.3. Step 3: demand modeling

The decision trees obtained in Step 2 provide two sets of decision rules, NewUtility() and RemanUtility(). The decision rules give estimates on customer utility that corresponds to a set of design attributes. NewUtility() gives the utility estimates in the current new market, and RemanUtility() gives the estimates in the future reman market.

Once customer utilities for a specific product and its competing products are given, it is possible to estimate the market share of each of the products. In this research, logit choice probability of the multinomial logit (MNL) model (Train, 2003) is used as shown in Equations (1) and (2), where $l$ and $m$ are the product choices available in the new and reman markets, respectively; $Y_{ij}$ is a vector of binary variables representing price related ($Y_{1j}$ for price of a new product, $Y_{2j}$ for price of a reman product) product attributes and their levels; $X_{ij}$ is a vector of binary variables representing component related product attributes and their levels; $MS^{new}$ and $MS^{reman}$ are the sizes of new and reman markets, respectively; $D^{new}(Y_{1j},X_{ij})$ and $D^{reman}(Y_{2j},X_{ij})$ are market demands for new and reman products, respectively.

$$D^{new}(Y_{1j}, X_{ij}) = \frac{\exp(NewUtiliy(Y_{1j}, X_{ij})}{\sum_1^l \exp(NewUtiliy_l(Y_{1j}, X_{ij}))} MS^{new} \tag{1}$$

$$D^{reman}(Y_{2j}, X_{ij}) = \frac{\exp(RemanUtiliy(Y_{2j}, X_{ij}))}{\sum_1^m \exp(RemanUtiliy_l(Y_{2j}, X_{ij}))} MS^{reman} \tag{2}$$

### 3.2. Optimal life cycle design

The optimal life cycle design model is the optimization engine for the PLCD. Table 3 shows the problem statement of the optimization model. With the aim to maximize the pre-life and end-of-life profits together, the model identifies the optimal product design as well as optimal production strategies at the pre-life and the end-of-life stages (i.e., the quantities and selling prices of new and reman products). The model assumes that the new products sold at time $t^{first}$ are all taken back for recovery after $h$ years at time $t^{eol}$. A certain percentage of the initial selling price, $\varepsilon \cdot P^{new}$, is paid for the take-back. It is also assumed that the returned end-of-life products are all recovered by either remanufacturing or recycling. Customer abuse and product reliability can affect the availability of remanufacturable products. Based on the product condition, only working products are allowed for remanufacturing. During the remanufacturing operation, no loss in yield or no scrap is assumed. Also, upgrades of parts are not considered. In other words, products

**Table 3**
Optimal life cycle design model.

| | |
|---|---|
| Objective function | *Maximize* (pre-life profit + end-of-life profit) |
| Decision Variables | - Product attributes<br>- Quantity of products to be manufactured and remanufactured |
| Constraints | - Design attributes uniqueness<br>- No excess fulfillment of products<br>- Take-back program |
| Given Info | - NewUtility () and RemanUtility ()<br>- Manufacturing, remanufacturing and recycling cost, and recycle profit<br>- Market size for new and reman products, and competing products<br>- Reusability or reliability of components |
| Assumptions | - Accumulated preference data are available<br>- Remanufacturing and recycling are possible recovery options<br>- No loss in yield in the recovery operation |
| Type of Problem | Mixed Integer Non-Linear Problem |

are remanufactured maintaining their initial design from the pre-life stage.

### 3.2.1. Objective function

The objective function of the model is given in Equation (3). It aims to maximize the total life cycle profit, i.e., sum of profits from the pre-life and end-of-life stages. Equation (4) formulates the total profit from the pre-life stage, i.e., the profit from making and selling $Q^{new}$ units of new products at the current time $t^{prelife}$. Equation (5) formulates the total profit from the end-of-life stage. It mainly consists of three parts: cost of taking back $Q^{takeback}$ units of end-of-life products, profit from remanufacturing $Q^{reman}$ units of end-of-life products, and profit from recycling $Q^{recycle}$ units of products. Since the end-of-life profit occurs at the future time $t^{eol}$, an annual interest rate $\alpha$ is applied to discount the value.

$$Maximize \quad f^{prelife} + f^{eol} \tag{3}$$

$$f^{prelife} = (P^{new} - C^{new})Q^{new} \tag{4}$$

$$f^{eol} = \frac{1}{(1+\alpha)^h}\left[\left(-C^{takeback} \cdot Q^{takeback} + (P^{reman} \cdot Q^{reman} - C^{reman} \cdot Q^{reman})\right) + \left(P^{recycle} - C^{recycle}\right)Q^{recycle}\right] \tag{5}$$

Equations (6) and (7) represent the prices of new and reman products as a function of the price related decision variable $Y_{ij}$. Equation (8) through (11) formulate the unit processing costs of manufacturing and recovery activities. In Equations (8) and (10), both manufacturing and remanufacturing costs are affected by binary decision variables, $X_{ij}$. If product attribute $i$ ($i \in I$) has the level of $j$ ($j \in J$), $Y_{ij}$ equals 1; otherwise, it equals 0. ü in Equation (9) denotes the take-back cost parameter and $C^{privacyprotection}$ represents the cost related to activities of privacy protection (e.g., data cleaning or scrubbing).

$$P^{new} = \sum_j P_{1j}^{new} \cdot Y_{1j} \tag{6}$$

$$P^{reman} = \sum_j P_{2j}^{reman} \cdot Y_{2j} \tag{7}$$

$$C^{new} = \sum_i \sum_j C_{ij}^{manufacturing} \cdot X_{ij} + C^{forwardlogistics} \tag{8}$$

$$C^{takeback} = \varepsilon \cdot P^{new} + C^{reverselogistics} + C^{sorting} + C^{privacyprotection} \tag{9}$$

$$C^{reman} = \sum_i \sum_j C_{ij}^{reconditioning} \cdot X_{ij} + C^{forwardlogistics} \tag{10}$$

$$C^{recycle} = C^{recycling} + C^{forwardlogistics} \tag{11}$$

### 3.2.2. Constraints

Equations (12) through (20) show the constraints of the model. Equation (12) imposes that each product attribute $i$ has an attribute level $j$. Equation (13) constrains the production quantity of new products, $Q^{new}$, in such a way that they are always less than or equal to the demand for them, $D^{new}(Y_{1j}, X_{ij})$. As described in the previous section, the demand is obtained by the decision-tree-based rules from the DTM. Similarly, Equation (14) constrains the production quantity of reman products, $Q^{reman}$.

$$\sum_j Y_{ij} = 1, \ \sum_j X_{ij} = 1, \ Y_{ij}, \ X_{ij} \in \{0, \ 1\} \tag{12}$$

$$Q^{new} \leq D^{new}(Y_{1j}, \ X_{ij}) \tag{13}$$

$$Q^{reman} \leq D^{reman}(Y_{2j}, \ X_{ij}) \tag{14}$$

Equation (15) formulates that available products from the new products sales at the first-life stage will be taken back for recovery at the end-of-life stage. $\rho$ denotes the take-back loss parameter due to the customer abuse. Equation (16) constrains that all the returned products are recovered either by remanufacturing or recycling.

$$Q^{takeback} = \rho \cdot Q^{new} \tag{15}$$

$$Q^{takeback} = Q^{reman} + Q^{recycle} \tag{16}$$

Equation (17) refrains $Q^{reman}$ from exceeding the available amount of remanufacturable products, $A(t^{eol})$. Equation (18) estimates $A(t^{eol})$, where it is determined by the multiplication of $Q^{takeback}$ and remanufacturability, $\delta(t^{eol})$, i.e., the probability that a product is still reusable and remanufacturable at the end-of-life stage. In Equation (19), $\delta(t^{eol})$ is defined as the multiplication of each part's reliability, $\gamma_j(t^{eol})$, at time $t^{eol}$. Because a part's reliability differs by design decisions, $\delta(t^{eol})$ is formulated as a function of $X_{ij}$. Finally, Equation (20) shows the variable conditions for production quantities.

$$Q^{reman} \leq A\left(t^{eol}\right) \tag{17}$$

$$A\left(t^{eol}\right) = Q^{new} \cdot \delta\left(t^{eol}\right) \tag{18}$$

$$\delta\left(t^{eol}\right) = \prod_i \left(\sum_j \gamma_j\left(t^{eol}\right) \cdot X_{ij}\right) \tag{19}$$

$$Q^{new}, Q^{reman} \in \text{nonnegative integer} \tag{20}$$

## 4. Illustrative example: smart-phone design

### 4.1. Overview

As the waste stream of discarded mobile phones grows rapidly, recovery of used phones has become an important issue in recent years. Mobile phones are known to have a relatively short life cycle,

**Table 4**
Assumptions about manufacturing and remanufacturing cost.

| | Manufacturing | | | | | | | | Remanufacturing | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Screen | | | Memory | | | Camera | | Screen | | | Memory | | | Camera | |
| | $X_{11}$ (2.8″) | $X_{12}$ (3.5″) | $X_{13}$ (5.3″) | $X_{21}$ (16 GB) | $X_{22}$ (32 GB) | $X_{23}$ (64 GB) | $X_{31}$ (8 MP) | $X_{32}$ (16 MP) | $X_{11}$ (2.8″) | $X_{12}$ (3.5″) | $X^{13}$ (5.3″) | $X_{21}$ (16 GB) | $X_{22}$ (32 GB) | $X_{23}$ (64 GB) | $X_{31}$ (8 MP) | $X_{32}$ (16 MP) |
| Cost ($) | 26 | 36 | 48 | 30 | 38 | 52 | 18 | 38 | 3.5 | 3.7 | 4 | 2.3 | 2.5 | 2.9 | 3 | 3.2 |

approximately 1.5 years (Cellular-Recycler, 2011). In 2009, the U.S. Environmental Protection Agency (EPA) estimated that Americans discard approximately 129 million mobile devices every year, of which only 8% are recycled properly (Environmental Protection Agency, 2011). This implies not only an environmental problem but also a missing profit opportunity. According to the EPA, “recycling one million cell phones can save enough energy to power more than 185 U.S. households with electricity for a year.” ReCellular, Inc is another testimony of profitable recovery. According to the Wall Street Journal (Pearce, 2008), “ReCellular resold 5.2 million mobile phones in 2010, up from 2.1 million five years earlier, and its revenue was $66 million.”

This section illustrates the PLCD framework with an example of smart-phone design. Suppose that there is an OEM smart-phone manufacturer that operates a one-year take-back program; they make and sell new products, and after one year, they take back the products for remanufacturing. For such take-back, it is assumed that the company returns 15% of the new-product price to the customer. To maximize the total profit from manufacturing and remanufacturing, the company aims to optimize their product design considering changing trends in the market. This section shows that the PLCD framework with the DTM can serve their needs effectively and demonstrates that the company can achieve greater profit by adopting the model. To be specific, there are five product attributes that the company wants to optimize: selling prices of new and reman products, screen size, memory size, and camera pixels. Depending on which attributes are chosen, the product would have different production costs and reliability, and different profits at the pre-life and end-of-life stages. Tables 4 and 5 present assumptions on production costs and part reliability for attribute choices.

### 4.2. Demand Trend Mining

To apply the DTM, two sets of customer preference data are required: one for the current new market and the other for the future reman market. The former is collected at a single time point $t^{prelife}$ and used for estimating market demand at time $t^{first}$. The latter, on the other hand, is collected over multiple time points from $t^1$ though $t^{prelife}$ and used for capturing future demand at $t^{eol}$. In this study, preference data were artificially generated. A total of 216 samples were simulated for each time point. The data for reman market was simulated as ten time-stamped data with six-month intervals; in other words, preference data reflecting market trends over the last five years were collected over ten time points, $t^1$ to $t^{10}$. Here, $t^{10}$ represents the current time $t^{prelife}$, i.e., $t^{10}t^{prelife}$. Since the time gaps between $t^{prelife}$ and $t^{first}$, and $t^{eol}$ and $t^{second}$ were very short for the simplicity, the historical data was used for the prediction of demands at $t^{12}$ with a one-year take-back program.

The data structure was the same as shown in Table 2. Each sample represented a specific combination of design attributes and the corresponding class variable (i.e., customer utility). As discussed in Section 3, all variables were defined as discrete variables. Table 2 shows design candidates of each variable.

In order to obtain decision rules, $NewUtiliy(Y_{1j},X_{ij})$ at $t^{prelife}$ ($=t^{10}$) and $RemanUtiliy(Y_{2j},X_{ij})$ at $t^{prelife+2}$ ($=t^{12}$), the C4.5 and PTM were applied to the new and reman market data, respectively. Weka 3.6.5 (Hall et al., 2009) and R 2.14.0 (R Development Core Team, 2008) were used for the decision tree induction and automatic time series forecasting. The resulting rules are given in Figs. 5 and 6. Each path in Figs. 5 and 6 represents a decision rule for a utility estimation. For example, in Fig. 5, one can estimate that if the selling price of a new product is $199, the camera resolution is 8-MP, and the memory is 16-GB, the screen size is 2.8-inch then the corresponding customer’s utility is 2 out of 4.

The decision rules in Figs. 5 and 6 allow estimation of the market share of a specific product. Suppose that the potential competing products are known as shown in Table 6. Then, the decision rules can calculate the utility of each competing product, which in turn enables to use Equations (1) and (2) for market share estimation.

### 4.3. Optimal life cycle design

Figs. 5 and 6 provide different rules for utility estimation. In other words, the market demands at the pre-life and end-of-life stages are different from each other. For example, a smart-phone with a $199 (for reman $199.5) price, 3.5-inch screen, 64-GB memory, and 8-MP camera would generate utility value 3 for new product and 2 for reman product. This implies that product design optimized based on the pre-life data only would not be optimal from the end-of-life perspective. To find an optimal product design, the optimal life cycle design model was applied.

In addition to the assumptions in Table 4 through Table 6, the following assumptions were made. The cost of reverse logistics, sorting, and data scrubbing is $2 in total and the cost of forward logistics is $1. The size of new market is 100,000 in terms of the total number of buyers, and the size of reman market is 50% of the new market. As shown in Table 5, the remanufacturability, or, reusability rate of a phone is less than 100%, which means that not all the new products can be remanufactured due to functional damages or poor conditions. The take-back loss parameter is one, and only working phones with good conditions would be remanufactured while the remainder is recycled. The recycling profit is $0.621 (USGS, 2006) and the recycling cost is $0.39 per cell phone (Bhuie et al., 2004). Lastly, to discount future profit from the end-of-life stage, an annual interest rate of 3% is assumed.

To solve the optimization problem, the Excel risk solver platform was used. Table 7 shows the optimization results (Column PLCD). To maximize the total life cycle profit, a smart-phone should be equipped with 2.8-inch screen, 64-GB memory, and 16-MP camera. The optimal selling price of the product is $399 at the pre-life stage; the optimal selling price of the remanufactured version is $149.5.

**Table 5**
Assumptions about part reliability after one year.

| Screen | | | Memory | | | Camera pixel | | |
|---|---|---|---|---|---|---|---|---|
| 2.8″ | $X_{11}$ | 0.95 | 16 GB | $X_{21}$ | 0.9 | 8 MP | $X_{31}$ | 0.92 |
| 3.5″ | $X_{12}$ | 0.92 | 32 GB | $X_{22}$ | 0.9 | 16 MP | $X_{32}$ | 0.88 |
| 5.3″ | $X_{13}$ | 0.88 | 64 GB | $X_{23}$ | 0.9 | | | |

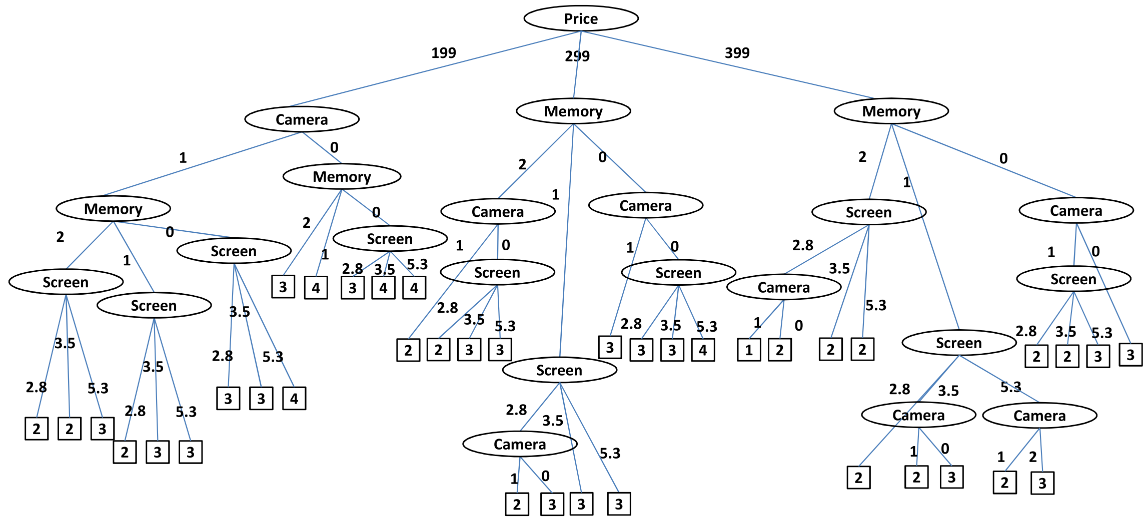


**Fig. 5.** Decision tree for new product at $t^{prelife}$ or $t^{10}$.

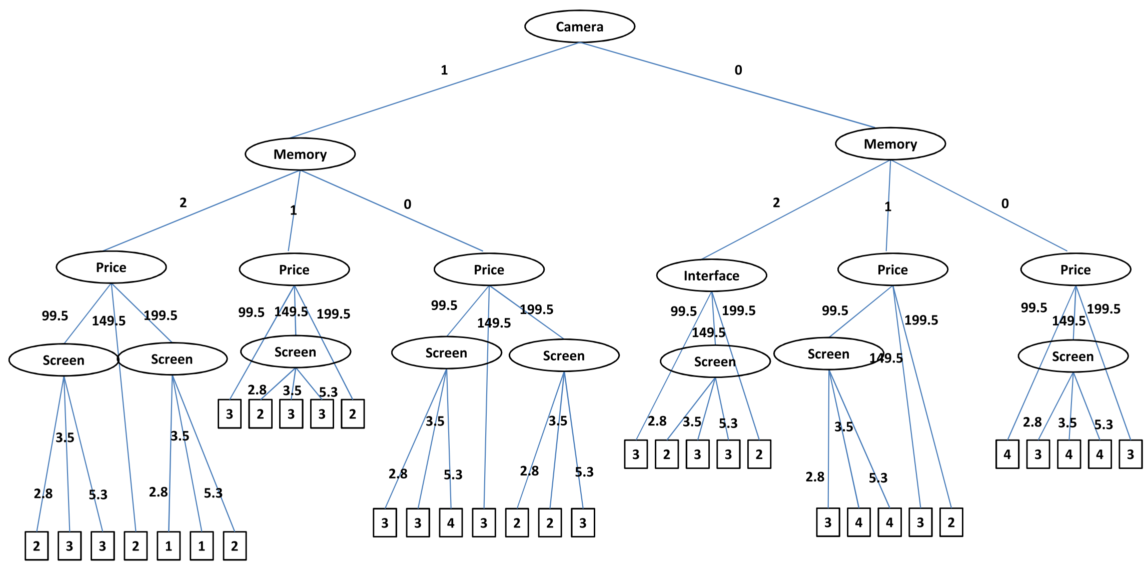


**Fig. 6.** Decision tree For Reman product at $t^{eol}$ or $t^{12}$s.

The optimal solution also provides optimal production and recovery strategies. The quantity of new products to produce should be 36,552 units; after one year, 26,722 units should be remanufactured, and the rest (9830 units) recycled. The maximum total profit results in $11,703,000 (in terms of the value at $t^{second}$).

### 4.4. Discussion

Many traditional design approaches have been focused on maximizing the profit from the pre-life stage only. The PLCD framework with the DTM algorithm is different from them in that it considers the entire life cycle of a product and maximizes the total profit from the life cycle. To demonstrate the benefit of the PLCD framework, this section compares the optimization result of the PLCD with those of traditional design approaches. To be specific, two traditional approaches are considered in this section, i.e., pre-life design without any end-of-life recovery and pre-life design with end-of-life recovery. Both approaches seek an optimal product design which maximizes the profit from the pre-life stage; they do not consider how their decision will affect the end-of-life stage. In the latter approach, however, the OEM conducts recovery at the end-of-life stage and tries to maximize the profit from recovery with additional optimization. The additional optimization means optimizing the production quantity and price of the reman product with pre-determined design attributes.

Table 7 shows the optimal design and the maximum profit from the traditional approaches. When the pre-life design is conducted, the product is optimized solely for the new market, and different attributes are chosen as the optimal: 3.5-inch screen, 32-GB memory, 16-MP camera. The maximum profit that can be achieved by this design is $10,490,000; if the company conducts recovery at the end-of-life stage (i.e., pre-life design with end-of-life recovery), the profit is increased by $67,000 to $10,556,000. Compared to the PLCD, the pre-life designs bring a greater profit at the pre-life stage. However, the benefit of the PLCD is revealed when the life cycle profit is considered. In Table 7, the profit from the PLCD is 10.9% higher than that of the pre-life design with end-of-life recovery.

**Table 6**
Assumptions about competitors information.

| | High spec product | | | | | Mid spec product | | | | | Low spec product | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Attributes | New price | Reman price | Screen | Memory | Camera | New price | Reman price | Screen | Memory | Camera | New price | Reman price | Screen | Memory | Camera |
| | $Y_{13}$ | $Y_{23}$ | $X_{13}$ | $X_{23}$ | $X_{32}$ | $Y_{12}$ | $Y_{22}$ | $X_{12}$ | $X_{22}$ | $X_{31}$ | $Y_{11}$ | $Y_{21}$ | $X_{11}$ | $X_{21}$ | $X_{31}$ |
| New utility | | | | | | | | | | | | | | | |
| | 3 | | | | | 2 | | | | | 2 | | | | |
| Reman utility | | | | | | | | | | | | | | | |
| | 3 | | | | | 3 | | | | | 2 | | | | |

**Table 7**
Comparative result between PLCD and pre-LIFE design.

| | | PLCD | Pre-life design | Pre-life design (+end-of-life later) |
|---|---|---|---|---|
| Total profit [$] | | 11,703,000 | 10,490,000 | 10,557,000 |
| Profit for pre-life [$] | | 10,344,000 | 10,490,000 | 10,490,000 |
| Profit for end-of-life [$] | | 1,359,000 | – | 67,000 |
| Product attributes | New product price [$] | 399 | 399 | 399 |
| | Reman product price [$] | 149.5 | – | 99.5 |
| | Screen Size [inch] | 2.8 | 3.5 | 3.5 |
| | Memory [GB] | 64 | 32 | 32 |
| | Camera Pixel [MP] | 16 | 16 | 16 |
| Quantity of reman product [EA] | | 26,722 | 0 | 26,722 |
| Quantity of recycled product [EA] | | 9830 | 0 | 9830 |
| New product utility | | 3 | 3 | 3 |
| Reman product utility | | 4 | – | 4 |

Previously, the size of the reman market was assumed to be half the size of the new market or $MS^{reman} = 0.5{*}MS^{new}$. However, as reported by Pearce (2008) and Cellular-Recycler (2011), the reman market is expected to grow more in the future. To see the effect of an increasing size of reman market and validate the outcome in Table 7, a sensitivity analysis is conducted. In Fig. 7, $\beta$ denotes the ratio of $MS^{reman}$ to $MS^{new}$. For both PLCD and pre-life design (with recovery) models, the sensitivity analysis examined how the maximum achievable profit changes as $\beta$ increases. A different selection of design attributes and consequent demands and amounts of remanufacturable products ($D^{reman}$ and $A$) are attributed for different gaps in the graph. If $\beta = 0$, there are no market or demands for the reman products, and no remanufacturing is conducted; if $\beta = 1$, the size of the reman market is the same as the new market. When $\beta = 0$, the optimizer will determine the optimal design attributes only from the pre-life stage for both models, which will generate the same design attributes with the total profit of $8,300,000. When $\beta > 0$, it is expected that the total profit from the PLCD framework is greater than that of the pre-life model except the case of selecting the same design attributes. The results in Fig. 7 show that both models choose all different designs when $\beta > 0$. When $\beta = 0.6$, the slops of the both models are changed since the upper bound is changed from $D^{reman}$ to $A$ (Equations (13) and (17)). When $\beta = 0.7$, both models select different designs from the previous ones. For $\beta = 0.8$ and $\beta = 0.9$, the upper bounds are changed again, and finally when $\beta = 1$, the optimal design is changed for the PLCD. In the illustration example, when $\beta = 0.9$, the profit difference is maximized. The results reaffirm that the PLCD framework with the DTM algorithm is always better than the traditional pre-life design, although the magnitude of the benefit changes depending upon $\beta$.

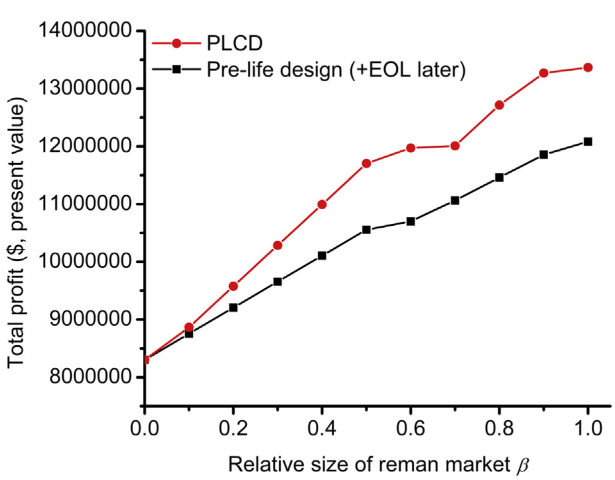


**Fig. 7.** Sensitivity analysis OF Reman market size ratio.

## 5. Conclusion and future work

This paper proposed a new demand modeling technique, data trend mining (DTM), for product design analytics. The first contribution is the development of the DTM algorithm. In order to capture hidden and upcoming trends of demand, the algorithm combines three different models: decision tree for large-scale data, discrete choice analysis for demand modeling, and automatic time series forecasting for trend analysis. The DTM algorithm dynamically reveals design attribute pattern that affects demands. The second contribution is the new design framework, predictive life cycle design (PLCD), which connects the DTM and data-driven product design. The optimization-based model enables a company to optimize its product design by considering the pre-life and end-of-life stages of a product simultaneously. The DTM model interacts with the optimization-based model to maximize the total profit of a product. The smart-phone case study demonstrated that there is a hidden source of opportunity for profit and the PLCD framework can help utilize this opportunity. Moreover, the sensitivity analysis reaffirmed that the life cycle design is more preferable than the traditional design method.

The current PLCD framework considers/optimizes two consecutive life cycles of a single product. In the future, the model can be extended to accommodate multiple life cycles and multiple products. The current DTM algorithm allows discrete attributes and class variables only, which should be extended to process continuous attributes and class variables. Also, in reality, it is possible that a product evolves with new attributes. Future work will also include incorporating emerging attributes into the DTM. Text mining (Tucker and Kim, 2011a; Rai, 2012) and sentiment mining (Stone and Choi, 2013) techniques in the domain of product design can be candidates for the management of dynamic attribute sets. On-line review data is a promising source that can provide not only customer preferences but also important emerging attributes.

**Acknowledgments**

The work presented in this paper is supported by the National Science Foundation under Award No. CMMI-0953021. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References


Bhuie, A., Ogunseitan, O., Saphores, J.D., Shapiro, A., 2004. Environmental and economic trade-offs in consumer electronic products recycling: a case study of cell phones and computers. In: Electronics and the Environment, 2004. Conference Record. 2004 IEEE International Symposium on, pp. 74–79.

Böttcher, M., Spott, M., Kruse, R., 2008. Predicting future decision trees from evolving data. In: Proceedings of ICDM '08, pp. 33–42.

Box, G.E.P., Jenkins, G., 1976. Time Series Analysis, Forecasting and Control. Holden-Day, Incorporated.

Cellular-Recycler, 2011. Sustainability within the Used Cellular Phone Industry. Report. Cellular Recycler. http://www.cellularrecycler.com/wp/wp-content/uploads/2011/01/CR-Sustainability-Report.pdf.

Charter, M., Gray, C., 2007. Remanufacturing and Product Design: Designing for the 7th Generation. Centre for Sustainable Design. Report.

Deng, L., Williams, E., Babbitt, C., 2009. Hybrid life cycle assessment of energy use in laptop computer manufacturing. In: Sustainable Systems and Technology, 2009. ISSST '09. IEEE International Symposium.

Duverlie, P., Castelain, J.M., 1999. Cost estimation during design step: parametric method versus case based reasoning method. Int. J. Adv. Manuf. Technol. 15, 895–906.

Environmental Protection Agency, 2011. Electronics Waste Management in the United States through 2009. EPA, U.S. Report EPA 530-R-11-002.

Fixson, S.K., 2004. Assessing product architecture costing: product life cycles, allocation rules, and cost models. In: ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2004), Salt Lake City, USA. DETC2004-57458.

Geurts, M.D., Ibrahim, I.B., 1975. Comparing the box-jenkins approach with the exponentially smoothed forecasting model application to Hawaii tourists. J. Mark. Res. 12, 182–188.

Gooijer, J.G.D., Hyndman, R.J., 2006. 25 years of time series forecasting. Int. J. Forecast. 22, 443–473.

Grissom, M., Belegundu, A., Rangaswamy, A., Koopmann, G., 2006. Conjoint-analysis-based multiattribute optimization: application in acoustical design. Struct. Multidiscip. Optim. 31, 8–16.

Hall, L.O., Chawla, N., Bowyer, K.W., 1998. Decision tree learning on very large data sets. In: In IEEE Conference on Systems, Man and Cybernetics, pp. 2579–2584.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The weka data mining software: an update. SIGKDD Explor. Newsl. 11, 10–18.

Harris, E., 2002. Information gain versus gain ratio: a study of split method biases. In: ISAIM.

Holt, R., Barnes, C., 2010. Towards an integrated approach to design for x: an agenda for decision-based dfx research. Res. Eng. Des. 21, 123–136.

Hucal, M., 2008. Product recycling creates multiple lives for caterpillar machines. Mag. Peoriamagazines. http://www.peoriamagazines.com/ibi/2008/sep/product-recycling-creates-multiple-lives-caterpillar-machines.

Hundal, M., 2001. Mechanical Life Cycle Handbook: Good Environmental Design and Manufacturing. In: Mechanical Engineering. Marcel Dekker.

Hyndman, R., Koehler, A., Ord, J.K., Snyder, R., 2008. Forecasting with Exponential Smoothing: the State Space Approach. Springer-Verlag, Berlin Heidelberg.

Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for r. J. Stat. Softw. 27 (3), 1–22.

Hyndman, R.J., Koehler, A.B., Snyder, R.D., Grose, S., 2002. A state space framework for automatic forecasting using exponential smoothing methods. Int. J. Forecast. 18, 439–454.

King, A., Miemczyk, J., Bufton, D., 2006. Photocopier remanufacturing at xerox UK a description of the process and consideration of future policy issues. In: Brissaud, D., Tichkiewitch, S., Zwolinski, P. (Eds.), Innovation in Life Cycle Engineering and Sustainable Development. Springer, Netherlands, pp. 173–186.

Kwak, M., 2012. green Profit Design For Lifecycle. Ph.D. thesis. University of Illinois at Urbana-Champaign.

Kwak, M., Kim, H.M., 2010. Evaluating end-of-life recovery profit by a simultaneous consideration of product design and recovery network design. J. Mech. Des. 132, 071001.

Kwak, M., Kim, H.M., 2011. Assessing product family design from an end-of-life perspective. Eng. Optim. 43, 233–255.

Kwak, M., Kim, H.M., 2013a. Design for lifecycle profit with a simultaneous consideration of initial manufacturing and end-of-life remanufacturing. Eng. Optim. 135.

Kwak, M., Kim, H.M., 2013b. Market positioning of remanufactured products with optimal planning for part upgrades. J. Mech. Des. 135.

Labrinidis, A., Jagadish, H.V., 2012. Challenges and opportunities with big data. Proc. VLDB Endow. 5, 2032–2033.

Lund, R.T., 1984. Remanufacturing: the Experience of the United States and Implications for Developing Countries. World Bank, Washington, D.C., U.S.A.

Lye, S.W., Lee, S.G., Khoo, M.K., 2001. A design methodology for the strategic assessment of a product's eco-efficiency. Int. J. Prod. Res. 39, 2453–2474.

Mangun, D., Thurston, D., 2002. Incorporating component reuse, remanufacture, and recycle into product portfolio design. Eng. Manag. IEEE Trans. 49, 479–490.

McGarry, K., 2005. A survey of interestingness measures for knowledge discovery. Knowl. Eng. Rev. 20, 39–61.

Moore, W.L., Louviere, J.J., Verma, R., 1999. Using conjoint analysis to help design product platforms. J. Prod. Innov. Manag. 16, 27–39.

Naylor, T.H., Seaks, T.G., Wichern, D.W., 1972. Box–Jenkins methods: an alternative to econometric models. Int. Stat. Rev./Rev. Int. Stat. 40, 123–137.

Newcomb, P.J., Bras, B., Rosen, D.W., 1998. Implications of modularity on product design for the life cycle. J. Mech. Des. 120, 483–490.

O'Shea, M., 2002. Design for environment in conceptual product design a decision model to reflect environmental issues of all life-cycle phases. J. Sustain. Prod. Des. 2, 11–28.

Parker, D., Butler, P., 2007. An Introduction to Remanufacturing. Report. Centre for Remanufacturing and Reuse.

Pearce, J.A., 2008. In with the old. Mag. Wall Str. J. http://online.wsj.com/news/articles/SB122427020019745211.

Quinlan, J., 1993. C4.5: Programs for Machine Learning. In: Morgan Kaufmann Series in Machine Learning. Morgan Kaufmann Publishers.

R Development Core Team, 2008. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Rai, R., 2012. Identifying key product attributes and their importance levels from online customer reviews. In: ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2012), Chicago, USA. DETC2012-70493.

Rose, C., Ishii, K., Stevels, A., 2002. Influencing design to improve product end-of-life stage. Res. Eng. Des. 13, 83–93.

Rose, C., Stevels, A., Ishii, K., 2000. A new approach to end-of-life design advisor (ELDA). In: Electronics and the Environment, 2000. ISEE 2000. Proceedings of the 2000 IEEE International Symposium on, pp. 99–104.

Seo, K., Park, J., Jang, D., Wallace, D., 2002. Approximate estimation of the product life cycle cost using artificial neural networks in conceptual design. Int. J. Adv. Manuf. Technol. 19, 461–471.

Stone, T., Choi, S.K., 2013. Extracting consumer preference from user-generated content sources using classification. In: ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2013), Portland, USA. DETC2013-13228.

Train, K., 2003. Discrete Choice Methods with Simulation. Discrete Choice Methods with Simulation. Cambridge University Press.

Tucker, C., Hoyle, C., Kim, H.M., Chen, W., 2009. A comparative study of data-intensive demand modeling techniques in relation to product design and development. In: Proceedings of the ASME Design Engineering Technical Conferences, San Diego, CA, USA. DETC2009-87049.

Tucker, C., Kim, H.M., 2008. Optimal product portfolio formulation by merging predictive data mining with multilevel optimization. J. Mech. Des. 130, 041103.

Tucker, C., Kim, H.M., 2011a. Predicting emerging product design trend by mining publicly available customer review data. In: Proceedings of International Conference on Engineering Design, Copenhagen, Denmark, pp. 43–52.

Tucker, C., Kim, H.M., 2011b. Trend mining for predictive product design. J. Mech. Des. 133, 111008.

Tucker, C.S., Kim, H.M., 2009. Data-driven decision tree classification for product portfolio design optimization. J. Comput. Inform. Sci. Eng. 9.

USGS, 2006. Recycled Cell Phones – a Treasure Trove of Valuable Metals. U.S. Geological Survey. Fact Sheet 2006-3097. Http://pubs.usgs.gov/fs/2006/3097/fs2006-3097.pdf.

Van Horn, D., Olewnik, A., Lewis, K., 2012. Design analytics: capturing, understanding and meeting customer needs using big data. In: ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2012), Chicago, USA. DETC2012-71038.

Wassenaar, H.J., Chen, W., 2003. An approach to decision-based design with discrete choice analysis for demand modeling. J. Mech. Des. 125, 490–497.

Wassenaar, H.J., Chen, W., Cheng, J., Sudjianto, A., 2005. Enhancing discrete choice demand modeling for decision-based design. J. Mech. Des. 127, 514–523.

Yu, Z., Haghighat, F., Fung, B.C., Yoshino, H., 2010. A decision tree method for building energy demand modeling. Energy Build. 42, 1637–1646.

Zhao, Y., Thurston, D., 2010. Integrating end-of-life and initial profit considerations in product life cycle design. In: ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2010), Montreal, Quebec, Canada. DETC2010-28830.