VIRAL PRODUCT DESIGN FOR SOCIAL NETWORK EFFECTS

A Dissertation
Presented to
The Academic Faculty

by

Feng Zhou

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mechanical Engineering

Georgia Institute of Technology
December 2014

VIRAL PRODUCT DESIGN FOR SOCIAL NETWORK EFFECTS

Approved by:

Dr. Roger J. Jiao, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Jonathan S. Colton
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Julie S. Linsey
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Ashok K. Goel
School of Interactive Computing
Georgia Institute of Technology

Dr. Heidi A. Hahn
Engineering Sciences Directorate
Los Alamos National Lab

Date Approved: November 7, 2014

ACKNOWLEDGMENTS

It would have been not possible for me to finish my doctoral dissertation without the help and support of the people around me, the excellent research environment provided by, and generous funding supported by Georgia Institute of Technology.

First and foremost, I would like to take this opportunity to express my sincerest gratitude and appreciation to my advisor Dr. Roger J. Jiao at Georgia Institute of Technology for his invaluable guidance, supervision, and advice for my Ph.D. study. I am greatly appreciated for Dr. Roger J. Jiao who helps shape my research ability and carries me on through difficult times with his insights and suggestions.

Special thanks go to my thesis reading committee, including Dr. Jonathan S. Colton (Mechanical Engineering), Dr. Julie S. Linsey (Mechanical Engineering), Dr. Ashok K. Goel (Interactive Computing), and Dr. Heidi A. Hahn (Los Alamos National Lab) for their precious time and suggestions to improve my dissertation in various ways.

I also would like to thank all the students in Room 264B MaRC building and staff at School of Mechanical Engineering at Gatech who always are ready to help me in times of need. Specifically, I would like to thank Yitao Liu, Dr. Yangjian Ji, Xiaoming Hu, Dr. Dong Yang, Dr. Dazhong Wu, Hui Xia, Thomas Stone, Glenda Johnson, Trudy Allen, and Dr. Mason Hollandbeck.

Last but not least, great thanks go to my family for always being there when I need them most, and for supporting me through all these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Chapter 3

σ	The expected number of adopters
S	A set of seed customers
$\|S\| = n$	The cardinality of S, i.e., the number of seeds in the set S is n
$V = \{v_1, \dots, v_i, \dots, v_N\}$	A set of N customers/social entities in the social network
$A = \{a_k\}_K$	A set of K (refined) product attributes
$A^V = A^{VP} \cup A^{VI}$	Viral attributes A^V is the union set of viral product attributes A^{VP} and viral influence attributes A^{VI}
L_k	The number of levels for the k-th product attribute a_k
$A^* = \{a_{kl}^* \| k = 1, \dots, K, l = 1, \dots, L_k\}$	A set of attribute levels, i.e., the value set of A
$X_j = \left(x_{j11}, \dots, x_{j1L_1}, \dots, x_{jK1}, \dots, x_{jKL_K}\right)$	Configuration of the j-th product P_j
$X = \left\{X_j \| j = 1, \dots, J\right\}$	A set of J products
$x_{jkl} \in \{0,1\}$	The value of l-th attribute level of the k-th attribute for P_j
$Y = \left(y_j \| j = 1, \dots, J\right)$	A vector indicating a particular choice of P_j
u_{ikl}	The utility of the l-th attribute level of k-th attribute perceived by the i-th customer
U_{ij}	The holistic product utility of P_j perceived by customer v_i
h_{ij}	Customer v_i's hurdle utility for the j-th product P_j
$z_{ij} \in \{0,1\}$	The value of customer v_i adopting P_j or not
J^+	The number of product variants offered in the market
$G = (V, E)$	A directional graph, where V is a set of vertices/nodes and E is a set of edges
I_i^t	The influence of social network effects of v_i from his/her active neighbors at time t
I_{ji}^t	The individual influence from active neighbor v_j to v_i at time t
ϑ_i	The adoption threshold of v_i
θ_i	The activation threshold of v_i
$\Pr\left(P_j \| i, t\right)$	The probability that v_i will adopt product P_j at time $t + 1$
S^D	Social network data

Chapter 4

p_v	Activation probability of social entity v

$b_{u,v}$	Influence weight from u to v
$p_{u,v}$	Influence probability of u on v
$u \in N_v^a$	u is one element of v's active neighbor set N_v^a

Chapter 5

$R^P = \left\{r_i^p\right\}_{i=1}^{I}$	A set of I review web pages
$F_r = \{r_i\}_{i=1}^{R}$	A file of R product reviews
$C^R = \{C_k^R\}_{k=1}^{M_R}$	A set of M_R refined use cases
$u_k = \{a_k, s_k, f_k\}_{k=1}^{K}$, $u_{kl} = \{a_{kl}^*, s_{kl}, f_{kl}\}_{l=1}^{L_k}$	a_k's or a_{kl}^*'s preference information
$s_{kl}, s_k = \{p^s, n^s\}$,	Sentiment orientation, p^s indicates positive and n^s negative
f_k, f_{kl}	Frequency of a_k, a_{kl}^*
$C_k^N = \left\{a_k, \{c_{kj}^n\}_{j=1}^{J_k}\right\}_{k=1}^{K}$	a_k's customer needs ($J_k$statements with the j-th c_{kj}^n)
w_k	The k-th word in ANEW
syn(w_k), ant(w_k)	Synonym set and antonym set of w_k
$s(w_k) = \{w_k^j\}_{j=1}^{n_w}$	A set formed by w_k, syn(w_k), and ant(w_k) with n_w elements
$A_p = [a_{ij}]_{n_w \times n_w}$	Propagation matrix
$s_k^0, s_k^{m_w}$	The initial score vector and the m_w-th score vector
m_w, λ_w	Parameters involved in the lexicon propagation algorithm
$s_{k\_n}^{m_w}$	The normalized score vector
N_{pos}, N_{neg}	The numbers of words with positive and negative valence in a sentence
$V_{ave}, A_{ave}, D_{ave}$	The average values of valence, arousal, and dominance in a sentence
V_{max}, V_{min}	The maximum and minimum valence in a sentence
N_n, N_a	The numbers of negation and adversative words in a sentence
$len(w_i, w_j)$	Path length between w_i and w_j
C_{HS}, k_{HS}	Constants involved in Hirst-St-Onge similarity function
d_{HS}	The number of direction changes in the path between w_i and w_j
D_{LC}	The maximum depth of the taxonomy in Leacock-Chodorow similarity function
$lso(w_i, w_j)$	The least common subsumer between w_i and w_j
$p_s(w)$	The probability of encountering an instance of a synonym of w

$pa_i = [u_1, \dots, u_{m_1}]$ A product attribute with m_1 terms

$fa_j = [w_1, \dots, w_{m_2}]$ A frequent product attribute with m_2 terms

$max_j\left(ws(u_i, w_j)\right)$ The maximum word similarity between u_i and w_j

N_{mfa} The number of matched attributes

N_{fa} The number of frequent attributes discovered by association rule mining

N_{mpa} The number of matched attributes in the user predefined set

$C^R =< C_1^R, \dots, C_{M_R}^R, Ind, R_d >$ Use case database, C_i^R is the i-th case, Ind is a case index model, and R_d is a domain knowledge model for case adaptation

C^e Extraordinary use case

C^{ijkl} Retrieved use case

C^{ijk*} Retrieved use case with any contextual event

C^{ij**} Retrieved use case with any contextual event and interaction environment

C^{i***} Retrieved case with any contextual event, interaction environment, and user type

$C_i = \left(c_1^i, c_2^i, \dots, c_{m_c}^i\right)$ The i-th use case with m_c characteristics

Chapter 6

$U = C(P, \mathbf{p})$ A holistic product utility of a product P obtained by a copula function C, where $\mathbf{p}$ is a choice probability vector

u_{kl} Individual part-worth utility of a_{kl}^*

$v_{kl} = v(a_{kl}^*)$ Subjective value function of a_{kl}^*

$a_{k,ref}^*$ The reference attribute level of product attribute a_k

α, β Parameters in [0,1], modulating the curvature of the subjective value function

λ $\lambda > 1$ specifies the degree of aversion to negative preferences

p_{kl} Choice probability of a_{kl}^*

η Positive scaling parameter for the logit model

η_j Sensitivity parameter for the *j*-th participant in the logistic choice rule

$P_{kl} = F(p_{kl})$ Cumulative probability of a_{kl}^*

$w(P_{kl}) = \pi(P_{kl}) - \pi\left(P_{k,l-1}\right)$ A decision weight of v_{kl}

$\pi(P_{kl})$ Weighting function of P_{kl}

$z = \delta$ or θ Curvature of the weighting function for positive or negative preferences

$d_{ji} = \{1, 0\}$ Result of j-th participant making the i-th choice of two products

N_d, M_p The number of decisions made for each participant, the number of participants

q_{ji} The logistic choice probability that accounts for choice inconsistencies

$\eta_j, \alpha_j, \beta_j, \lambda_j, \delta_j,$ and θ_j Individual parameters involved in the customer preference model

Φ Standard normal cumulative distribution function

$z_j^\alpha = \Phi^{-1}(\alpha_j)$ Probitized individual parameter of α_j

$N(\mu^\alpha, (\sigma^\alpha)^2)$ A normal distribution with mean μ^α and standard deviation σ^α

$LN(\mu^\lambda, (\sigma^\lambda)^2)$A lognormal distribution with respective location, scale parameters μ^λ, σ^λ

$\kappa_j = [\alpha_j, \beta_j, \delta_j, \theta_j, \eta_j, \lambda_j]$ 6-tuple individual parameters

$\kappa = [\kappa_1, \ldots, \kappa_M]$ $6 \cdot M$ tuple for M participants

$\mathrm{K} = [\mu^\alpha, \sigma^\alpha, \mu^\beta, \sigma^\beta, \mu^\theta, \sigma^\theta, \mu^\delta, \sigma^\delta, \mu^\eta, \sigma^\eta, \mu^\lambda, \sigma^\lambda]$ Group-level parameters

$\Theta = [\vartheta_1, \ldots, \vartheta_n] = [\kappa, \mathrm{K}]$ Parameters to estimate in the customer preference model

u_{kl_k} Normalized value of individual part-worth utility u_{kl}

$\varphi(\xi_k) = \frac{1-\exp(\varsigma\xi_k)}{1-\exp(-\varsigma)}$ Generating function in the copula function, $\varsigma \neq 0$

$c_1 = 1/\left(1 - \varphi^{-1}(\prod_{k=1}^{K} \varphi(\xi_k))\right), c_2 = 1 - c_1, \xi_k$ Parameters in the copula function

Chapter 7

$G = (V, E)$ A directed graph with a set of social entities V and a set of edges E

$G = (V, E, L)$ A directed graph with an influence function $L: E \to [0,1]$

Δ_j Price of product P_j

I_{ji}^t, x_{ji}^t Normalized and original strength of the social tie from v_j to v_i at time t

E_{ji}^t, d_{ji}^t Normalized and original Euclidean distance from v_j to v_i at time t

$v_i = \left(c_1^i, c_2^i, \ldots, c_{m_e}^i\right)$ The i-th social entity with m_e characteristics

S_{ji}^t Normalized entity similarity between v_j and v_i at time t

p_{ji}^t Influence probability from social entity v_j to social entity v_i at time t

p_i^t Influence probability from v_i's active neighbors to v_i at time t

$N_i^{a,t}$ A set of active neighbors of v_i at time t

$R_h, R_m, \bar{R}_{v_k \in N_i^a}$ Highest, median, and mean ratings by v_i's active neighbors at time t

ρ_{ij}, μ_{ij} Probabilities of promoting product P_j when v_i is in the states of adopt and tattle

I_i^T, S_i^T, E_i^T Interaction strength, entity similarity, and structural equivalence from $v_i{}'$s active neighbors at current time T

$I^S = (\Pi, F)$ An information system, such that $\forall \mathbf{f} \in F: \Pi \rightarrow F^*$, where Π is a non-empty finite set called the universe, F is a non-empty feature set, for any feature vector $\mathbf{f} = \{f_l\}_L \in F$

$\mathbf{d} = \{d_h\}_H \in D$ A decision vector with H decision variables and D is a decision set

$\Gamma = (F \cup D, I)$ A decision table, where $F \cup D$ is the universe of inference I

$I_i^* \sim (F_i^* \rightarrow D_c^*)$ Inference relationship from the features F_i^* to the decision D_c^*

$F_i^* = \{f_{il}^*\}_L$ The value set of the feature vector **f**

$D_c^* = \{d_c^*\}_C$ The value set of decision vector **d**

$\Phi = \{\phi_k\}_K \subset \Gamma$ A reduct, which is a proper subset of decision table Γ

$\phi_k = (f_i^{\phi}, d_i^{\phi})$ An indiscernibility relation, in which, for objects $x \in \Pi$ and $y \in \Pi$, a pair $(x, y) \in \Pi \times \Pi$ belongs to Φ

$f_q^{\phi}(z), d^{\phi}(z)$ q feature instances and one decision instance about an object $z \in \Pi$

tst A test sample

$SupportSet(r)$ A set of training examples matching rule r

$MatchRules(tst, d_c^*)$ A subset of minimal rules applicable to tst and decision d_c^*

$|\cdot|$ The cardinality of a set '·'

V_N^T A set of non-adopter up to the current time T

tem A template in decomposition tree

$\Gamma(tem), \Gamma(\overline{tem})$ Decision sub-tables with samples matching template tem or not

$O(N^2)$ A growth rate on the order N^2

Chapter 8

$\mathbf{X} \in \mathbf{R}^s$ A s-dimensional design variable

$\mathbf{Z} \in \mathbf{R}^t$ A t-dimensional design variable

$F(\mathbf{X}, \mathbf{Z}), f(\mathbf{X}, \mathbf{Z})$ Leader and follower

$G(\mathbf{X}, \mathbf{Z}) \leq 0, g(\mathbf{X}, \mathbf{Z} \leq 0)$ Leader's and follower's respective constraints

$\Omega = \{(\mathbf{X}, \mathbf{Z}): G(\mathbf{X}, \mathbf{Z}) \leq 0, g(\mathbf{X}, \mathbf{Z}) \leq 0, \mathbf{X} \in \mathbf{R}^s, \mathbf{Z} \in \mathbf{R}^t\}$ Constraint region

$\mathrm{U} = \{\mathbf{X} \in \mathbf{R}^s: \exists \mathbf{Z} \in \mathbf{R}^t, \text{such that } (\mathbf{X}, \mathbf{Z}) \in \Omega\}$ Feasible set for **X**

$\mathrm{R}(\mathbf{X}) = \{\mathbf{Z} \in \mathbf{R}^{\boldsymbol{t}}: \mathbf{Z} \in \operatorname{argmin}\{f(\mathbf{X}, \bar{\mathbf{Z}}): g(\mathbf{X}, \bar{\mathbf{Z}}) \leq \mathbf{0}\}\}$ Follower's rational reaction set

$\text{IR} = \{(\mathbf{X}, \mathbf{Z}) : (\mathbf{X}, \mathbf{Z}) \in \Omega, \mathbf{Z} \in \text{R}(\mathbf{X})\}$ Inducible region

$\phi(X, Y)$ Shared surplus function

$\sigma(S, X, Y)$ Adoption maximization function

C_j A cost function of product P_j

ϱ A constant indicating the average dollar cost per variation of process capabilities

LSL^T, μ_j^T, σ_j^T The lower specification limit, the mean, and the standard deviation of the estimated cycle time for product P_j

$g_{jk}(X) \geq 0$ Technical constraints of product j for the k attribute domains

$F(X, \Psi(X))$ Single-level parametric optimization function

$Y = \Psi(X)$ A unique response function of X

$L_m(q^{m-1}) = [a_{i,j}]_{m \times (m-1)}$ A standard orthogonal array, where L denotes a Latin square, $m = q^k$ is the number of experiment runs, $k > 1$ is a positive integer, $m - 1$ is the number of columns in the orthogonal array

$\{o_i | i = 1, \dots, n_o\}$ A set of n_o observations

$c_j = X_j = \{x_{j1}, x_{j2}, \dots, x_{jK}\}$ A chromosome with attribute levels, $x_{jk} \in \{1, 2, 3\}$

$fit(c_j)$ Fitness function

$p\left(c_j, \Psi(c_j)\right)$ Penalty function

C_p Penalty constant

$r = \{r_1, \dots, r_K\}$ A string of uniformly generated random numbers

M_p The number of initial population in HTGA

p_c, p_m Crossover probability and mutation probability

E_{fl} Effect of factor f at level l

$I_{max}, I_{min}, C_{min}$ The maximum, minimum iteration numbers and minimum convergence number

ε_{ij} A random error term for each segment-product pair

$D^l = (V, P, T)$ A log of past product adoption for all the customers

$\langle i, P_j, t \rangle$ A tuple indicates customer v_i adopted the product P_j at time t

$L(D^l; \Theta)$ Likelihood of the data maximization problem given the model parameters Θ

$\delta(i, P_j)$ A delta function, it is 1 if v_i adopted P_j, and 0 otherwise

AdpMaxVA	Adoption maximization with viral attributes
ANEW	Affective Norms for English Words
ANOVA	Analysis of Variance
CBR	Case-Based Reasoning
CELF	Cost-Effective Lazy Forward
CPT	Cumulative Prospect Theory
GA	Genetic Algorithm
HTGA	Hybrid Taguchi-Genetic Algorithm
IDEF0	Icam DEFinition for Function Modeling
IFE	In-flight Entertainment
InfMax	Influence Maximization
KFHD	Kindle Fire HD
LT model	Linear Threshold model
LTH model	Linear Threshold-Hurdle model
MCMC	Markov Chain Monte Carlo
RBF	Radial Basis Function
SNR	Signal-to-Noise Ratio
SoC	Share-of-Choice
SVMs	Support Vector Machines
UML	Unified Modeling Language
MOOCs	Massive Open Online Courses

SUMMARY

Recent advances in social media have profound technical and economic implications for innovative design. This research is motivated to investigate social network effects on product design with a focus on the interface of engineering design, viral marketing, and social computing. This dissertation envisions a new paradigm of design, called *viral product design for social network effects*. The research problem is formulated as identification of both an optimal set of product configurations and an optimal set of seed customers so as to maximize product adoption via online social networks through equilibrium solutions to marketing-engineering coordination. Fundamental issues are investigated and a technical framework is proposed with integrated decision-based design methods. Results of case studies demonstrate that the proposed research is able to bridge the gaps between the domains of engineering design and viral marketing by incorporating social network effects.

The proposed work is geared towards new design theory and decision models by integrating peer influence of social networks, which shed light on understanding the social aspect of design. The dissertation reveals the fundamental issues underlying *viral product design*, including the identification of viral attributes, customer preference modeling incorporating subjective experiences, the dynamics of the diffusion mechanism of online social networks, formulation of adoption maximization, and coordination between the marketing and engineering domains. In order to tackle the fundamental issues, a technical framework of viral product design for social network effects is proposed. Accordingly, mathematical and computational models are developed within the framework to support 1) latent customer needs elicitation for viral product attributes extraction, 2) customer preference modeling and quantification for product choice decision making, 3) social network modeling for product adoption prediction, and 4) viral product design evaluation by adoption maximization. These coherent models along the technical framework lay the theoretical foundation of this research, as described below.

First, in order to extract potential viral product attributes, latent customer needs elicitation is emphasized. This is because latent customer needs can delight customers

unexpectedly, and thus lead to potential product adoption to a large extent. We propose to elicit latent customer needs by use case analogical reasoning from sentiment analysis of online product reviews. A case study of Kindle Fire HD tablets shows the potential and feasibility of the proposed method. The extracted product attributes and attribute levels provide the choice set of viral product attributes.

Second, based on the extracted product attributes, a customer preference model based on cumulative prospect theory is presented, accommodating subjective experiences in the product choice decision making process. Moreover, a hierarchical Bayesian model with Markov chain Monte Carlo is used to estimate parameters involved in the model. Based on the case study of aircraft cabin interior design, the model parameters under different experimental conditions show systematic influence of subjective experiences in choice decision making. Furthermore, a copula structure is used to construct a holistic product utility, showing customers' overall preferences to a product. This measure is crucial to product choice decision making in the context of social networks.

Third, in order to predict product adoption incorporating peer influence of social networks, a linear threshold-hurdle model is proposed. It overcomes multiple drawbacks of traditional diffusion models by modeling activation thresholds, influence probability, adoption spread, holistic utility of the product, and hurdle utility of a customer in a holistic fashion. A case study of Kindle Fire HD tablets demonstrates both the predictive power of the proposed model and interesting results about customers' adoption behavior. This model paves the way for product adoption maximization in large social networks.

Fourth, in order to coordinate between marketing-engineering concerns, I formulate a bi-level game theoretic optimization model for viral product design evaluation, in which the leader maximizes product adoption, while the follower optimizes product line performance. Through social network effects in terms of viral product attributes and viral influence attributes, the expected number of product adopters and the expected shared surplus, resulting from the identified product configurations and seed customers, are proved to be larger than those obtained from existing practice of viral marketing and product line design respectively, based on the case study of Kindle Fire HD tablets. Thus, the proposed paradigm of design extends the traditional boundaries among domains of engineering design, viral marketing, and social computing.

CHAPTER 1

INTRODUCTION

This chapter provides an overview of the background knowledge leading to this research. Based on the discussion of research motivation, the research is identified as *viral product design for social network effects*, which suggests itself as an important strategy to achieve maximum product adoption, while considering both customer satisfaction and engineering concerns in the context of online social networks. Accordingly, research objectives and scopes are defined, along with an outline of a technical roadmap for this research.

1.1 Research Motivation

Recent advances in social media that have created an amazing fabric of connected social networks have profound technical and economic implications for design and innovation research. These social networks not only make modern products (e.g., iPhone and iPad) "participate" in social media to ignite conversations, elicit emotions, and engender loyalty, but also make customers more interconnected and informed when they make product choice decisions. Therefore, these social networks play a fundamental role as media for the spread of information, ideas, and influence among their social entities (Kempe et al., 2003). Understanding such a role in product design and marketing is critical to the success of the product. Some products would spread quickly to a large population (e.g., a new mobile phone among college students), whereas others would die out (a new weed spray in a village) in the social network (see Rogers, 2003). Three most important aspects are identified for such phenomena.

(1) Social networks: The first aspect arises from social networks (e.g., Twitter, Facebook, Google+, and review sections of shopping websites). Three important effects have been identified in the social networks, including word-of-mouth effects, imitation effects, and network effects (Dou et al., 2013). First, the word-of-mouth effects about a product, a service, or a brand often speed up the process of information diffusion and awareness (Dichter, 1966). The words of mouth from the social entities in the social network not only help catalyze the awareness of a product, but also influence other people to adopt the product (Narayan et al., 2011). For example, Trusov et al. (2010) find that, on average, about one-fifth of a user's friends actually influence his or her activity level (e.g., purchase behavior) on social networking sites. Second, increased adoption of specific products or services often leads to herding behavior, especially when there is homophily (McPherson et al., 2001) in the social network and when the number of adopters has surpassed the threshold (Granovetter, 1978). This process is often depicted as the result of imitation effects (Dou et al., 2013). Third, a network effect (or network externality) refers to the effect that one user of a product or a service has on the value of that product to other users (Shapiro and Varian, 1999). When a product (e.g., a smartphone) is prone to a positive network effect, the more people adopt it in the social network, the more valuable it is to each user. It thus increases the tendency of product adoption.

These effects can be understood as *social network effects* and have been studied in different domains with different types of behaviors, such as peer influence, neighborhood effect, conformity and contagion (Iyengar et al., 2011). In this research, I call the influence caused by social network effects as *peer influence* of social networks. Many researchers have realized the importance of social networking sites as a tool to influence potential customers to adopt a product. For example, Aral and Walker (2011) design an experiment to generate econometrically identifiable social influence and social contagion effects about a product in a large Facebook network. Kempe et al. (2003) introduce two basic diffusion models, namely the independent cascade model and the linear threshold model, to understand the difussion mechansim in the social network. Meanwhile, approximation algorithms are proposed to maximize product influence in the context of social networks. Since then, multiple difffusion models (e.g., Bhagat et al., 2012; Goyal et al., 2010; Lu and Lakshmanan, 2012) and algorithms (e.g., Chen et al., 2009; Goyal et al., 2011) have been proposed to understand product adoption under peer influence and to maximize product adoption, respectively.

(2) Viral marketing: Viral marketing makes use of viral influence attributes to promote a product usually in the context of online social networks. These viral influence attributes include personalized referrals, passive broadcasting, tagging, commenting on likes and dislikes, and inputting product reviews, and so on. For example, Aral and

Walker (2011) study the effect of personalized referrals and passive broadcasting on product adoption in Facebook, and find that despite the fact that personalized referrals are more effective in influencing product adoption, passive-broadcasting reaches more people and often causes a larger overall product adoption. By making use of these viral influence attributes, viral marketing is actually highly intertwined with social networks. First, viral marketing can take various forms to transmit product-related information, such as images, text, messages, emails, flash, and video clips, via social network services. Second, these transmissions also have various forms, including pass-along based, incentive based, and trendy based, and so on. In such a way, the information can be spread through self-replicating viral processes, which quickly increase product awareness among different types of potential customers (Howard, 2005).

Another important viral influence attributes are the social network users who are potential customers of a product. And by identifying an optimal set of seed customers so that based on their influence (i.e., social network effects), the expected number of adopters of the product can be maximized in the social networks, i.e., influence maximization (InfMax) (Kempe et al., 2003). This problem has been widely studied in viral marketing (e.g., Chen et al., 2009; Goyal et al., 2011). However, the interplay between product design and viral marketing in the context of online social networks is surprisingly still largely unexplored. Furthermore, one of the prominent limitations of the prevailing methods in viral marketing is that only seed customers by making use of their peer influence are considered in the InfMax problem. The underlying assumption is that the holistic product utility (as a way to measure customer preferences) is considered to be the same for all the customers and is thus ignored.

(3) Engineering design: Research has shown that another aspect that causes a product to be viral in the social network are the product attributes. For example, Berger and Milkman (2010) find that awe-inspiring news stories that are practically useful, surprising, positive, or affect-laden are more likely to make it into the New York Times "most e-mailed" articles list. Heath et al. (2001) show that disgusting urban legends are more likely to be shared. In this sense, many product attributes make it possible to diffuse in the social networks. Aral and Walker (2011) differentiate viral product attributes and viral influence attributes. The former is about the content of the product while the latter is related to how the product is shared and interacted among the social entities within a social network.

One of the important engineering methods to identify potential viral product attributes is product portfolio planning (also known as product line design), which has two stages (Li and Azarm, 2002). The first stage is product portfolio identification, which aims to identify a set of product attributes and attribute levels. The second stage is product portfolio evaluation, which can be formulated as the share-of-choice (SoC) problem (Camm et al., 2006). It aims to select a near-optimal mix of product variants configured by different product attribute levels to offer in the target market to increase the market share or to maximize its product adoption (Jiao et al., 2007b). In this sense, the common practice is to incorporate potential viral product attributes by understanding customer preferences. Customer preferences are often derived from conjoint analysis-based part-worth utilities, based on which the product variant with the maximum weighted sum of part-worth utilities is expected to maximize product adoption (e.g., Camm et al., 2006; Wang et al., 2009).

However, one of the weaknesses of the SoC problem is that it does not consider engineering concerns, such as costs and manufacturability involved in different product variants (Du et al., 2014). It assumes that any combination of product attributes observed by customers in a conjoint analysis study can be designed by engineers post hoc. This is often questionable for some moderately complex product in which engineering tradeoffs cannot be balanced between customer preferences and engineering constraints (e.g., costs and manufacturability) without considering marketing-engineering coordination (Michalek et al., 2011). In this situation, the SoC problem leads to non-optimal products with regard to product line performance and company profits. Another limitation in the SoC problem is that it ignores the social interactions between customers, or the social context in the product adoption process. It thus cannot make use of social network effects.

Based on the discussions about the InfMax problem that incorporates peer influence for adoption maximization and the SoC problem that incorporates potential viral product attributes for adoption maximization, it is imperative to combine these two questions in a systematic frame. Hence, I propose *viral product design for social network* *effects*, i.e., the process of explicitly engineering both viral product attributes and viral influence attributes into products so that they are more likely to be shared amongst peers and to generate peer-to-peer influence in its adoption process. It not only aims to maximize product adoption in the context of online social networks from the perspective of viral marketing, but also considers customer preferences and engineering concerns from the perspective of engineering design. In this sense, a product or product line design that has the best compromise, i.e., the equilibrium solution, between the marketing domain and the engineering domain is considered to be optimal.

1.2 Research Objective

The primary objective of this research is to formulate a systematic framework of viral product design for social network effects so that both product adoption and product line performance can be jointly optimized. Accordingly, it is decomposed into several sub-objectives that are to answer the following key research issues:

1) How to formulate viral product design systematically;
2) What are the fundamental issues of viral product design;
3) How to solve these fundamental issues systematically;
4) What are the key factors and the operational mechanism that make products viral;
5) How to test principles of viral product design under different scenarios with rigorous, transparent, and replicable methodologies.

Towards this end, corresponding core research tasks are proposed:

1) Examine relationships of the InfMax problem and the SoC problem and formulate viral product design as a bi-level game theoretic optimization problem, in particular

   a. Identify the contribution of viral product attributes and viral influence attributes to product adoption in the paradigm of viral product design;
   b. Analyze interactive relationships between viral product attributes and viral influence attributes and integrate them together for viral product design;
   c. Formulate viral product design systematically with a mathematical model.

2) Propose a technical framework to develop rigorous methodologies for viral product design that bridges the gaps between the marketing domain and the engineering domain. The framework consists of four consecutive and iterative design steps, i.e., latent customer needs elicitation for viral product attributes extraction, customer preference modeling and quantification for product choice decision making, social network modeling for product adoption prediction, and viral product design evaluation by adoption maximization. Corresponding research tasks have been conducted as follows:

For latent customer needs elicitation:

a. Predict sentiment orientation/polarity of online product reviews;
b. Extract product attributes, attribute levels, and use cases from online product reviews;
c. Mining customer preferences from sentiment analysis of product reviews;
d. Transforming explicit customer needs into latent customer needs by use case analogical reasoning from sentiment analysis of online product reviews, and the corresponding product attributes that satisfy latent customer needs become one part of design space for viral product design.

For customer preference modeling and quantification:

a. Develop a systematic model for customer preference quantification of different product configurations;
b. Investigate subjective influence on product choice decision making;
c. Validate the model with behavioral experiments by parameter shaping in the model;
d. Aggregate individual part-worth utilities considering their interdependence for a holistic product utility.

For social network modeling:

a. Represent a social network with an appropriate graph;
b. Develop a product diffusion and adoption model incorporating peer influence;
c. Predict product adoption based on the proposed diffusion and adoption model;
d. Validate the model based on a real-world data set of a case study.

For viral product design evaluation:

a. Propose appropriate evaluation measures that consider both marketing and engineering goals;
b. Formulate viral product design with a bi-level game theoretic optimization problem systematically;
c. Propose a solution strategy to solve the bi-level game theoretic optimization problem;
d. Evaluate viral product design based on the results of a case study.

1.3 Research Scope

Viral product design for social network effects is proposed as a new paradigm to approach product design incorporating peer influence. Based on the research objective and research issues, it attempts to bridge the gaps between viral marketing and engineering design in the context of online social networks. As shown in Figure 1.1, first, for latent customer needs elicitation, it involves applying social computing, i.e., sentiment analysis of online product reviews, for engineering design purposes, i.e., product attributes extraction and latent customer needs elicitation. Second, customer preference modeling and quantification are rooted in engineering design. In this research, from a behavioral science point of view, I especially investigate the influence of human affective elements in the product choice decision making process for customer preference modeling. Third, innovation diffusion and communication have been widely studied in social sciences. In this research, social network modeling for product adoption prediction is studied in the context of an online social network. Specifically, its diffusion mechanism is modeled, and customer's decision making process of product adoption is predicted with a data mining method for the purpose of viral marketing. Finally, viral product design incorporating viral product attributes and viral influence attributes is formulated as a bi-level game theoretic optimization problem for the purpose of evaluation. The upper-level model aims to maximize product adoption as a marketing goal and the lower-level model aims to optimize product line performance as an engineering goal. As a way to validate the proposed viral product design for social network effects, four case studies emphasizing different aspects in engineering design and viral marketing are conducted. Thus, the proposed research spans over the intersection and interaction of customers, products, and social contexts by integrating fundamental principles from multiple disciplines across domains of engineering design, viral marketing, and social computing.

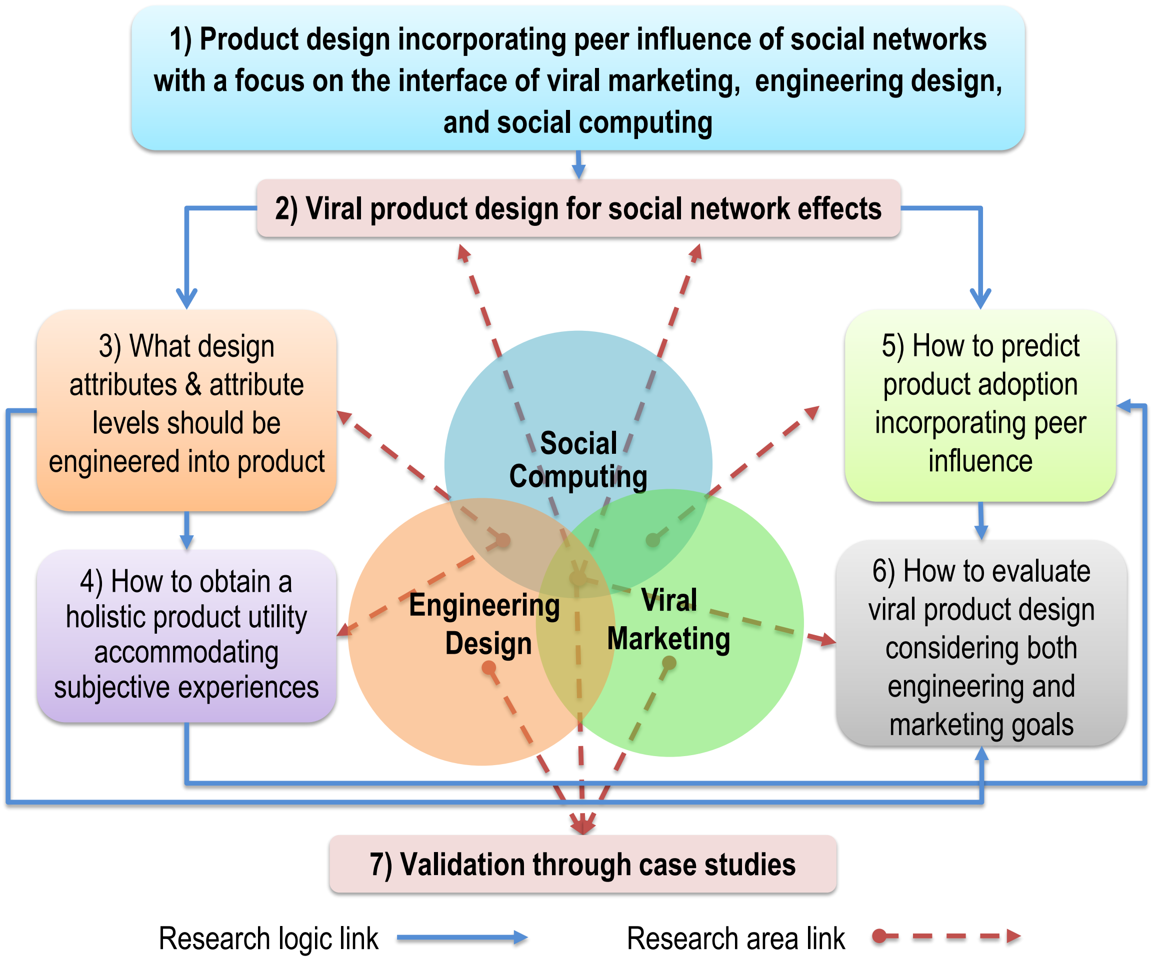


Figure 1.1 Research scope and research methodology

1.4 Organization of This Dissertation

Figure 1.2 presents the technical roadmap of this dissertation, including motivation & significance, problem formulation, technical approach, methodology & solution, and validation & application.

The motivation and significance are discussed in Chapters 1 and 2. Chapter 1 discusses the general background and a holistic view of this research. Chapter 2 provides a comprehensive review of various topics related to this research.

Chapter 3 formulates the key problems of this research. It presents the fundamental issues underlying product design incorporating peer influence of social networks. These fundamental issues help provoke insights into how to solve them systematically.

Chapter 4 proposes a technical framework along with a four-step viral product design process in order to solve the fundamental issues identified in Chapter 3. Their challenges and respective solution strategies are also discussed. Chapters 5, 6, 7, and 8 emphasize one step a time with the corresponding methodology, technical approach, and case study or experiment design for validation, respectively.

Chapter 5 conducts use case analogical reasoning from sentiment analysis of online product reviews for latent customer needs elicitation. Support vector machines are used to predict sentiment orientation of online product reviews based on an affective lexicon list. Association rule mining is used to extract product attributes and use cases. Subsequently, case-based analogical reasoning is used for latent customer needs elicitation. The product attributes that can satisfy latent customer needs are helpful in product diffusion in the social network.

Chapter 6 reports the development of customer preference modeling and quantification based on cumulative prospect theory accommodating subjective experiences. Choice decision making experiments are designed with regard to product attribute levels, and customers' affective states, cognitive tendency, and risk attitudes are manipulated in the experiments. Systematical patterns are found in terms of parameter shaping in the cumulative prospect-based customer preference model. Finally, a copulas-based method is used to aggregate individual part-worth utility functions for a holistic product utility to capture their interdependence.

Chapter 7 is devoted to social network modeling for product adoption prediction based on a linear threshold-hurdle model. We first identify three important operational factors underlying social network effects in order to quantify peer influence of social networks. Individual customers' hurdle utilities are compared with customers' perceived holistic utilities of the product in the adoption process. Finally, a data mining method named rough set theory is used to predict product adoption according to the linear threshold-hurdle model.

Chapter 8 focuses on the bi-level game theoretic optimization formulation for viral product design evaluation by investigating the interplay between product design and viral marketing. The product adoption maximization problem is modeled as the leader (i.e., the upper-level model) and the product portfolio optimization problem is modeled as a follower (i.e., the lower-level model). The interaction and coupling of these two optimization problems are addressed with a coordinate-wise optimization strategy iteration by iteration, in which adoption maximization and product line performance optimization are tackled by an improved greedy algorithm and a hybrid Taguchi genetic algorithm, respectively.

From Chapter 5 to Chapter 8, case studies of Kindle Fire HD tablets with different emphases and an experiment study of aircraft cabin interior design are used to illustrate the respective proposed methodologies. Each chapter focuses on different aspects of viral product design for social network effects as discussed above. The last chapter, Chapter 9, summarizes the achievements in addressing the research objectives and issues. A critical assessment is given to highlight the limitations and possible improvements of this research, along with recommendations for future work.

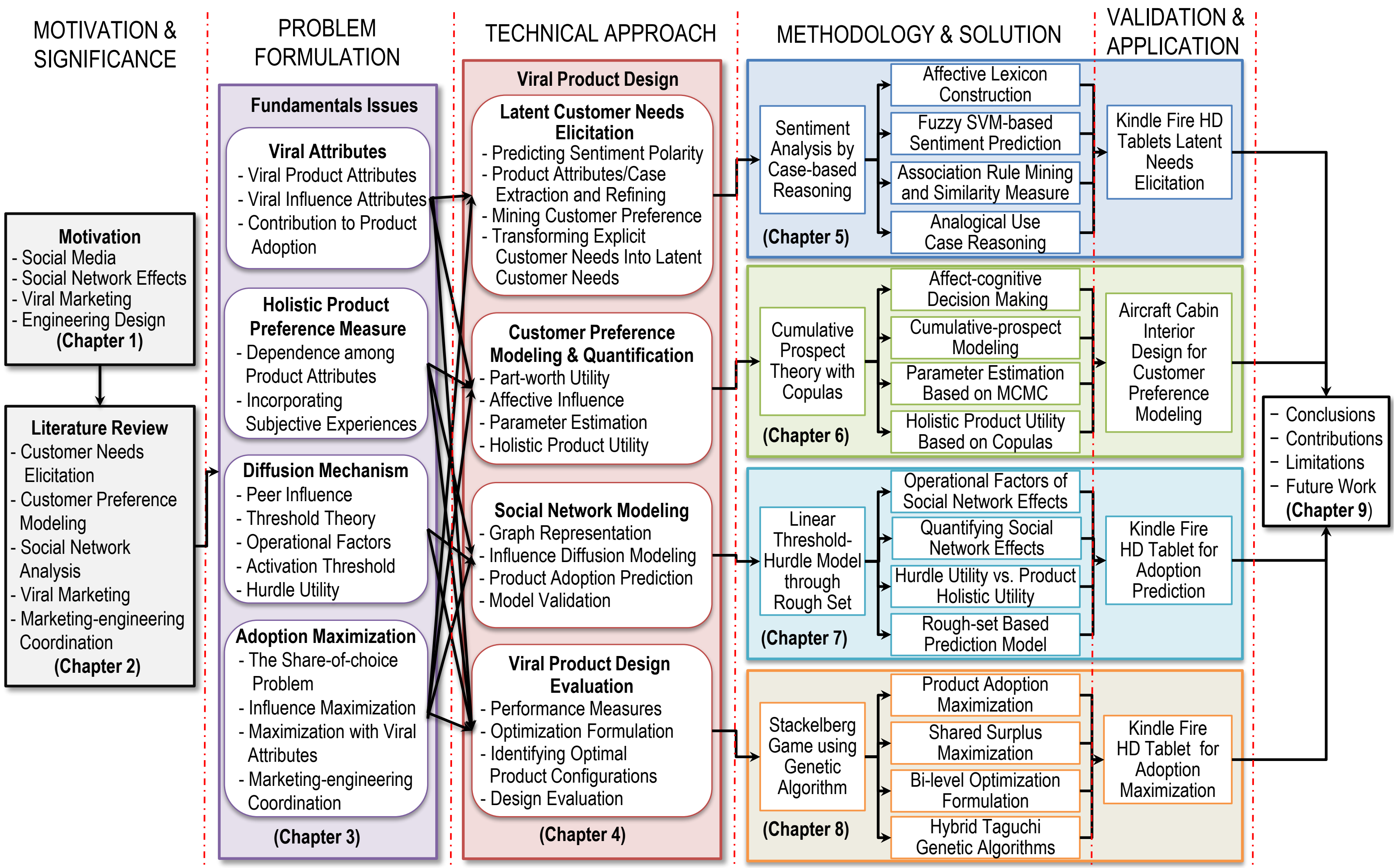


Figure 1.2 Organization of this dissertation

CHAPTER 2

LITERATURE REVIEW

This chapter is dedicated to the state-of-the-art review of viral product design for social network effects. Based on the research scope in Chapter 1, I will review various topics related to this research in engineering design, viral marketing, and social computing. A framework of reference will be given first that elaborates the topic relationships among different research domains. The limitations of the reviewed topics will also be discussed, which lead to the proposed methods in different chapters in the following.

2.1 Framework of Reference

As shown in Figure 2.1, this research mainly spans three domains, including engineering design, viral marketing, and social computing. *Engineering design* aims to build a product with a specified performance goal. It usually has a multi-step process, including task clarification, conceptual design, embodiment design, and detail design (Pahl et al., 2007). In this research, most related efforts in engineering design focus on viral product attributes with a goal to maximize customer satisfaction and/or expected number of product adopters, such as customer needs elicitation and customer preference modeling. Based on the customer needs elicited, especially latent customer needs, customer desired product attributes can be identified and are used as the choice set for viral product attributes. Another stream of work is modeling customer preferences in order to predict their choice decisions among product attribute levels or product variants, such as discrete choice analysis. Therefore, I mainly focus on the front end of engineering design, including customer needs elicitation and customer preference modeling. Various topics as shown in Figure 2.1 will be reviewed below.

In the domain of marketing, the related work is *viral marketing*. It refers to the marketing that use social media related services to produce increases in brand awareness and product sales through self-replicating viral processes (Howard, 2005). In this research, I mainly make use of viral influence attributes, which are related to features in the social network, such as personal referrals, automated broadcast notifications, tags, like, and dislikes, as well as seed customers. In this aspect, two important topics are product diffusion models and influence maximization (InfMax). The former focuses on how to model the product diffusion process in the social network, based on which algorithms are developed to maximize the influence of the product by identifying a small number of seed customers, i.e., InfMax.

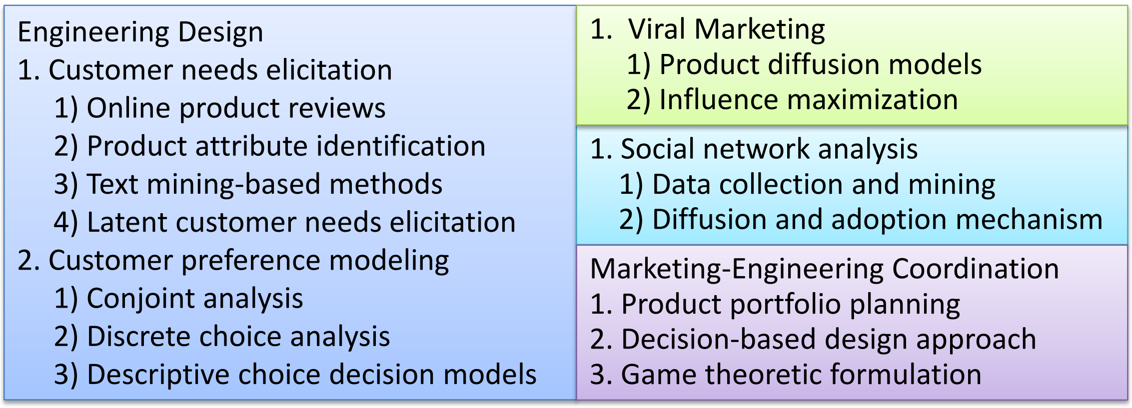


Figure 2.1 Various topics reviewed and their corresponding domains

However, in order to model the product diffusion process, it is important to understand the mechanism of product diffusion and adoption. This is one of the core questions in social network analysis in the domain of *social computing* (or computational social science) (Lazer et al., 2009), which is concerned with the interaction between social behavior and computational systems. Due to the large scale of social networks, it is necessary to apply efficient computational methods to collect a large amount of data, based on which data mining methods can be used to extract the underlying patterns.

Another important related topic is marketing-engineering coordination, in which both engineering concerns and marketing problems are considered in order to obtain an optimal result, such as product portfolio planning (which needs to consider marketing-engineering coordination, but often ignores it), decision-based design approaches, and game theoretic formulation. Note that some of the topics actually span across multiple domains, such as InfMax, which is studied both in social computing and viral marketing.

2.2 Customer Needs Elicitation

Understanding and fulfilling each individual customer's needs is of vital importance to customer satisfaction and product success. It thus is the basic requirement for a product to be viral and adopted by possibly a large number of customers. In order to elicit customer needs, it is important to identify customer desired product attributes. How customers use the product and what attributes they interact with indicate what customers like and dislike (Chen et al., 2013d). Recently, text mining techniques are used to identify product attributes (Putthividhya and Hu, 2011) and to elicit customer preferences and needs (Liu, 2010; Liu and Zhang, 2012) from online product reviews. However, one of the difficulties is to elicit latent customer needs, which are often the unexpected delighter for customers. The product attributes that can satisfy latent customer needs thus become the candidates of viral product attributes. Therefore, it is imperative to extract product attributes and reason different use cases extracted from online product reviews for latent customer needs elicitation. Compared with traditional methods, I mainly review those by making use of online user-generated content, such as online product reviews.

2.2.1 User-Generated Content

The exploding reach of the Web and the prevalence of social networking sites have created a huge amount of online user-generated content, which is considered as public data and is easily available. Among many, online product reviews have an important role for retailers, customers, and designers (Decker and Trusov, 2010). These user-generated product reviews often describe product performance in terms of different product attributes in different use situations from the user's perspective (Bickart and Schindler, 2001). Such use cases provide a specific channel for customer needs elicitation. Furthermore, studies have shown that unlike product information generated by sellers, which often emphasizes the performance of a product based on its technical specifications, users evaluate their purchased products based on their individual preferences and product performance in their specific situations (Chen and Xie, 2008). Therefore, being able to analyze user-generated online product reviews provides marketing and engineering a competitive advantage by understanding customer preferences, identifying customer needs, and predicting product demands, and so on.

The current literature mainly focuses on its function on product choice decision making and predicting product demands. For example, strong positive correlations have been found between positive ratings and growth of product sales (Clemons et al., 2006), and the quality of the reviews as measured by helpfulness votes also positively influences sales (Chen et al., 2008). Similarly, Ghose and Ipeirotis (2011) analyze the review features, such as subjectivity, informativeness, readability, and linguistic correctness, on product sales and their perceived usefulness. Miao et al. (2009) apply sentiment analysis of online product reviews to generate a ranking mechanism with temporal opinion quality and relevance to facilitate product choice decision making for customers. Archak et al. (2007) incorporate textual content of product reviews to learn customer preferences of product features and propose a model to estimate how the textual content can be used to predict consumers' product choices and demands. As pointed out by (Lee, 2007), however, prior analysis of online product reviews appears to have overlooked the role in customer needs elicitation. These online reviews can be a new approach to assessing rapidly changing customer needs.

2.2.2 Product Attribute Extraction

Product attribute extraction is one of the tasks in customer needs elicitation. Recently, sentiment analysis of online product reviews can be used to extract product attributes. Sentiment analysis is the computational study of opinions, sentiments, and emotions expressed in online texts (Liu, 2010). With regard to online product reviews, positive comments and negative comments on different product attributes can be identified using sentiment analysis. For example, Ghani et al. (2006) propose a text mining method to extract attribute and value pairs (i.e., attribute levels) from textual product descriptions online. Putthividhya and Hu (2011) combine supervised named entity recognition with bootstrapping to identify product attributes with a high precision of 90.33%. These two methods, however, are supervised and semi-supervised, which need a laborious manual labeling process for training. Hu and Liu (2004a, b) apply association rule mining to extract frequent items in the product reviews as candidates for product attributes. The advantage of this method is that it is unsupervised and thus no training process is needed. However, the number of attributes discovered for each product is over 100, and thus apparently unmanageable for consumer goods, despite the fact that a pruning strategy is applied, including attribute pruning and compactness pruning. Raju et al. (2009) propose an unsupervised approach to extract product attributes from

Amazon.com with 92% of precision, but only with 62% of recall. The low recall is often due to the fact that nouns and noun phrases are always considered as product attribute candidates. Zhao et al. (2010) then utilize syntactic structures of product attributes to improve the recall of product attribute extraction. However, another issue in these methods is attribute redundancy. For example, image quality and photo quality are often considered as two product attributes.

From the perspective of product design, more information about the extracted product attributes is often needed. For example, Wassenaar et al. (2005) map customer desires to design attributes related to engineering analyses, based on which a discrete choice demand model is proposed to assess product profits under the frame of decision-based design (Wassenaar and Chen, 2003). Tucker and colleagues (Tuarob and Tucker, 2013; Tucker and Kim, 2011) propose a statistical trend detection technique to help identify product attributes and classify them as standard, nonstandard, or obsolete based on the customer preference information. Park and Lee (2011) propose a framework based on text mining, co-word analysis, and decision tree to identify product attributes. Then clustering analysis is conducted about the comments on these attributes to identify customer groups. Archak et al. (2007) emphasize the weight customers placed on each product attributes and how these weights affect the product demand. Rai (2012) proposes a text mining method to partition online customer reviews into individual product attributes, and three importance measures are presented to rank identified attributes. However, automatic attribute level extraction is still an open question. In this research, in Chapter 5, I propose to apply an unsupervised method, i.e., association rule mining, to extract product attributes, based on which similarity measures based on WordNet (Miller, 1995) are used to refine extracted attributes by reducing attribute redundancy. Based on the preference information from customers' comments, I can further rank these attributes in terms of their frequency. Then attribute-level pairs are identified with the designer's scrutiny.

2.2.3 Text Mining-Based Methods

Based on the extracted attributes, the next step is to summarize customer needs. Many traditional qualitative methods, such as interviews, focus groups, questionnaires, self-reports, and observations (Crandall et al., 2006; Kinnear and Taylor, 1995; Wilson, 2014) have been proposed. More information can be referred to (Jiao and Chen, 2006; Zhou et al., 2013). Recently, sentiment analysis based on text mining can be used to elicit customer needs from online product reviews. It can classify product reviews into positive or negative opinions at the document level or the sentence level (Liu, 2010; Liu and Zhang, 2012). Generally, both supervised and unsupervised methods have been proposed. One of the unsupervised methods is to apply affective lexicons, which make use of semantic features (polarity tags and semantic orientations) in the product review. For example, Hu and Liu (2004a) form a list of lexicon seeds with known polarities, and then are expanded using WordNet (Miller, 1995) through synonym and antonym links to predict review orientation. Ding et al. (2008) improve Hu and Liu's method by including linguistic rules. Titov and McDonald (2008) propose a multi-aspect sentiment model, in which latent Dirichlet allocation is used to build topics representative of product attributes. These methods are domain independent and unsupervised, which make the whole system easy to implement and maintain. However, compared with supervised learning methods, their prediction accuracy can be limited, since no training process is involved.

For supervised methods, Jin et al. (2009) integrate multiple linguistic features, including part-of-speech tags, tag patterns, and lexicons under a frame of hidden Markov models, which are used to extract product attributes and classify sentiments at the same time with good performance. Their computation-intensive training process is alleviated by a bootstrapping process. Chen et al. (2012) propose models based on conditional random fields with similar linguistic features. Their model is able to outperform that of Jin et al. (2009) due to the fact that conditional random fields can overcome the limitations of hidden Markov models, i.e., inability to represent distributed hidden states and complex interactions among labels. However, one main limitation of these supervised methods is that they are often not product-independent and re-training models for other products is often expensive. In Chapter 5, I combine both affective lexicons and a supervised method, i.e., fuzzy support vector machines (SVMs) to elicit customer needs. First, affective lexicons are domain-independent, and can be applied to any online opinionated reviews for different kinds of products (e.g., movies, services, and physical products). In order to further improve the prediction accuracy with machine learning methods, a perfect candidate of binary classifiers is SVMs, which excel at separating categories by a clear gap that is as wide as possible. SVMs have been widely applied and proved to be effective in different areas in pattern recognition (e.g., Shao and Lunettab, 2012; Zhou et al., 2007). The results of sentiment analysis pave the way for latent customer needs with use case analogical reasoning.

2.2.4 Latent Customer Needs Elicitation

Many companies realize the difficulties in identifying latent customer needs, and therefore they focus on explicit customer needs and stress the importance to elicit the customers' real needs (Hanski et al., 2014). However, latent customer needs are critical to product innovations and are the value added, such as emotional needs (Zhou et al., 2013). For example, Yanagisawa and Murakami (2007) propose a shape generation system in which users can exchange design solutions of other users, which externalizes their latent emotional sensitivity. Zhou et al. (2010) apply rough set and K-optimal rule discovery to identify hidden relations between emotional needs and design elements for a truck cab interior design. From the design process perspective, Carlgren (2013) investigates how design competencies can contribute to the identification of latent needs. Yang (2013) proposes an analytical method to identify latent needs based on a refined Kano model through the provision of attractive and innovative product attributes.

Another important method to elicit latent customer needs is to interview lead users who experience needs still unknown to the public (Hannukainen and Hölttä-Otto, 2006). However, it is often difficult to identify lead users in large numbers. One solution to this issue is to transfer ordinary users into lead users by changing use cases. For example, by putting users in situations so that they are hard to see or hear, the use case makes users interact with products in extraordinary ways (Hannukainen and Hölttä-Otto, 2006). In this sense, these situational disabled users become lead users and interviewing them helps elicit latent customer needs. Lin and Seepersad (2007) propose a similar idea to transform ordinary customers into empathic lead users by creating extraordinary use cases (e.g., mitts on users' hands in the dark to simulate dusk on a cold day) for tent assembly. Experimental results from empathic lead user interviews significantly increase the effectiveness in latent customer needs discovery. Chen et al. (2013a) propose usage context-based design, and by changing the use context, latent customer needs are uncovered to show their choice decision making process. Unlike these studies, in Chapter 5, I capitalize on online product reviews without the interview process. First, the proposed method identifies extraordinary use cases by sentiment analysis, based on which case-based reasoning is applied for use case analogical reasoning to elicit latent needs by reusing and adapting ordinary use cases. It automates the latent needs elicitation process to a great degree.

2.3 Customer Preference Modeling

Viral product design involves two aspects, namely how to identify viral product and influence attributes and how to maximize product influence in the social network, so that their joint efforts contribute to maximal product adoption. Traditional methods are concerned with eliciting customer preferences and predicting customer choices by identifying product attributes with highest likelihood to be chosen by customers, such as conjoint analysis (Green and Srinivasan, 1978, 1990) and discrete choice analysis (Ben-Akiva and Lerman, 1985; Train, 2003).

2.3.1 Conjoint Analysis

Product adoption is often related to the twin questions, i.e., understanding customer preferences and predicting product choices. One of the most popular methods to estimate and quantify customer preferences is probably conjoint analysis (Green and Srinivasan, 1978, 1990). It models customer preferences by using part-worth utilities of different products' attribute levels. In such a framework, customers can decide which aspects of products are important, compare the products on each important aspect, and decide which one to choose (Louviere, 1988). Therefore, it is useful for marketers to optimize product configurations, simulate market response to new or modified offerings, and diagnose competitive strength (North and de Vos, 2002). For example, Jiao et al. (Jiao et al., 2007a) apply conjoint analysis to measure different affective product configurations of truck cabs based on ratings on a 9-point Likert scale. Green et al. (1992) report that conjoint analysis has been used to design a wide variety of products, including cell phones, computers, shipping services, pharmaceuticals, cable TV, and hotels, and so on.

However, there is little consensus about how to identify product attributes with assigned levels, and thus practice varies widely (Louviere et al., 2010). The combinations of attribute levels are often generated by orthogonal fractional arrays, but still customers are required to tediously rank order 20 to 40 products in terms of preferences, which resulting in an ordinal preference, i.e., how much one product is preferred over another (Hauser and Shugan, 1980). In addition, Louviere et al. (2010) state that conjoint analysis is concerned with the behavior of sets of numbers in response to factorial manipulations of attribute levels rather than human preference behavior.

2.3.2 Discrete Choice Analysis

In the customer preference modeling, one of the important tasks is to predict which product the customer will choose based on the customer preference information. One technique often used for customer choice modeling is discrete choice analysis originated from mathematical psychology (Ben-Akiva and Lerman, 1985; Train, 2003). It is a probabilistic choice modeling technique using a latent utility function to model a choice. The latent utility is formed by a systematic component and a random component. The former captures the utility of product attributes, explaining differences in choice alternatives, while the latter explains unidentified factors that impact choices (Louviere et al., 2010). Therefore, the utility function representing customer preferences is inherently stochastic due to the random component, which leads to a probability that a customer will choose one alternative among a choice set (Train, 2003). One major advantage of discrete choice analysis is the ability to capture the heterogeneity of customer preferences observed by customer characteristics (e.g., demographic factors, socio-economic attributes, anthropometric variables) (Chen et al., 2013d). It can be further used to estimate product demand by using sample enumeration for example (Ben-Akiva and Lerman, 1985). Different types of discrete choice models have been proposed to capture heterogeneity in customer preferences in different applications. For example, Chen et al. (2013d) propose analytic techniques, especially different discrete choice models to capture heterogeneous customer preferences so as to predict customer choices as a function of engineering design decisions and a targeted market population. These techniques include multinomial logit models (Hausman and McFadden, 1984), nested logit models (Garrow and Koppelman, 2004), and mixed logit models (McFadden and Train, 2000). For more information about discrete choice analysis, please refer to (Chen et al., 2013d).

2.3.3 Descriptive Decision Choice Models

Another important kind of preference choice model is descriptive decision model. Unlike normative models that are concerned with identifying the optimal choice to choose by assuming full rationality of the decision maker, descriptive decision models study how humans make decisions as it is, rather than as it should be, such as prospect theory (Kahneman and Tversky, 1979), cumulative prospect theory (CPT) (Tversky and Kahneman, 1992), disappointment theory (Bell, 1985), decision field theory (Mellers, 2000), and priority heuristics (Brandstätter et al., 2006). Among them, prospect theory and CPT perhaps are the most influential and successful decision models. For example, Chow et al. (2010) apply prospect theory to estimate a discrete choice model for choosing a high-occupancy-vehicle lane or not based on the current and perceived speeds. Compared with the binary logit model, they discover that the prospect theoretic-based model is more consistent with empirical data. Likewise, Xu et al. (2011) apply CPT to model a traveler's route choice decision making, and find that the results are more coherent with the experiment data than those obtained from the route choice model based on expected utility theory.

From the decision making point of view, many descriptive decision making models focus on cognitive aspects and heuristics in human judgments, but still ignore the role of emotion in human decision making (Brandstätter et al., 2006). Such a single cognitive perspective is not optimal for analyzing human decisions towards customer preferences, in which their affective states experienced at the time of decision making often influence their perceptions and choices (Ahn, 2010). Furthermore, recent affective neuroscience and psychological studies have reported that human affect and emotional experience play a significant role in human learning and decision making. For example, Ahn and Picard (2005) propose an affective-cognitive decision framework for learning and decision making. Zhang and Liu (2009) present a navigation system based on affective-cognitive learning and decision making, which speeds up the learning process and improves the capability of autonomous navigation. Power et al. (2011) emphasize the role of emotion in combination with health behavior models to provide a framework for conceptualizing patient decisions. Bracha and Brown (2012) propose an affective decision making model of choice under risk and uncertainty, and they posit that observed choices are the result of a rational and emotional interaction process with a Nash equilibrium. Penolazzi et al. (2012) study the role of impulsivity and reward sensitivity in affective and deliberative risky decision making, and suggest that personality traits differentially alter decision-making behavior due to interactions with the decision-making context. In order to incorporate affective factors in descriptive decision models for customer preference modeling, in Chapter 6, I propose a CPT-based decision making model to describe customer preference-based product choices, in which three different affective states and two types of products (i.e., affect-rich and affect-poor) are manipulated to show their influence on product choice decision making.

2.4 Social Network Analysis

Those product attributes that can satisfy customer needs and are preferred by customers, to some extent, make the product to be viral. However, the online social network context makes it possible for the product can be diffused and adopted by a potentially maximal number of customers through peer influence of social networks (e.g., viral influence attributes). In this regard, it is important to collect a large amount of user-generated data on social networks, based on which patterns can be extracted to understand social network effects, viral influence attributes, and the diffusion and adoption mechanism.

2.4.1 Data Collection and Mining of Social Networks

With the emergence of social media (e.g., Twitter and Facebook), things that are said on these social networks are considered as public data. The essential tool for analyzing a large amount of data collected is data mining, which has the capacity to collect and analyze data at a scale that may reveal patterns of individual or group behavior. Netvizz is data collection and extraction software that can help researchers export data in a standard file format from different types of Facebook social networking services (Rieder, 2013). Gelphi (https://gephi.github.io/) and NodeXL (http://nodexl.codeplex.com/) are two popular tools for social network data analysis, especially in visualization for complex systems and dynamic, hierarchical graphs.

Based on the data collected, data mining techniques can expand the researchers' capability of understanding new phenomena prevalent in social media, and in turn improve business intelligence to provide better services or develop innovative opportunities. Domingos and Richardson (2001) apply Markov random fields to study the network value of customers (i.e., the expected profit from sales to other customers who may be influenced to buy). They (2002) further propose an optimal viral marketing plan via mining patterns from knowledge-sharing websites, such as epinions.com. Aral et al. (2009) investigate the social contagion through a dataset of 262,985 Facebook pages and their associated fans. Chen et al. (2010) study the InfMax problem, in which the mined social networks are as large as those with the number of nodes ranging from 15,000 to 262,000 and the number of edges ranging from 31,000 to 1.2 million. Tuarob and Tucker (2013) mine a large scale of data from Twitter to successfully predict demand and longevity of various smart phones. Gundecha and Liu (2012) discuss the representative research problems of mining social media. In terms of the individual level and the societal level, Lazer et al. (2009) introduce the structure and content of relationships in social networks over an extended period of time.

2.4.2 Diffusion and Adoption Mechanism

The evolution and the dynamics of a product diffusion and adoption process under social network effects are very complicated. Many researchers have investigated this emerging research in the social aspect of product design and marketing. One immediate question is how to experimentally prove the existence of social network effects. Aral and Walker (2011) provide experimental evidence in a large Facebook network that viral influence attributes can generate econometrically identifiable social influence and social contagion effects; and further that firms can create word-of-mouth effects and social contagion by designing viral influence attributes into their products and marketing campaigns. Kumar et al. (2013) propose and implement a methodology to measure social media return on investment and a customer's words of mouth, and show positive relationships between social media and growth in sales, return on investment, and positive words of mouth. Centola (2010) experimentally shows that individual adoption is much more likely when participants receive social reinforcement from multiple neighbors in the social network. Such behavior spreads further and faster across clustered-lattice networks than across corresponding random networks.

In sociology, one influential milestone that explains the social diffusion and contagion is the threshold theory proposed by Granovetter and his colleagues (Granovetter, 1978; Granovetter and Soong, 1988). It assumes that individuals' behavior depends on the number of other individuals already engaged in that behavior (termed as a behavioral threshold). The threshold model is used to explain riots, residential segregation, and the spiral of silence, as well as other applications (including social contagion and peer influence). According to the threshold theory, Valente (1996) explicitly describes the social network thresholds in the diffusion of innovations. He creates a social network threshold model based on adopter categories (innovators, early adopter, early majority, late majority, and laggards) for both individual social networks and the entire social system. The model is used to (a) explain the definition of behavioral contagion; (b) predict the diffusion pattern of innovations; and (c) identify opinion leaders and followers in the social network. Consistent with this idea, Bearden and Etzel (1982) term the people who significantly influence others as the reference group, for which three distinct psychological processes are identified in terms of social influence from the public opinion point of view, namely, internalization, identification, and compliance (Kelman, 1961). From the diffusion and adoption process point of view, the book *Diffusion of Innovations* by Rogers (2003) describes the diffusion as a process by which an innovation is diffused through certain channels over time among the members of a social system. Four elements influence the spread process, i.e., the innovation, communication channels, time, and a social system.

2.5 Viral Marketing

2.5.1 Product Diffusion Models

In viral marketing, InfMax is one of the most frequently studied problems. Two important aspects are identified, including diffusion modeling and algorithm development. Among many, two classic types of diffusion models are proposed, including independent cascade models and linear threshold models (Kempe et al., 2003). Given a social network G modeled as a directed graph, each node is either active (i.e., product adopter) or inactive. Each node's tendency to become active increases monotonically as more of its neighbors becomes active. The adoption process unfolds in discrete steps. Assuming v is an inactive social entity (i.e., non-adopter) at step t, and the probability of v becoming active (i.e., adopters) depends on the influence of v's neighbors who are adopters at step t. In the linear threshold model, each node v has a threshold ϑ_v uniformly distributed between 0 and 1. When the joint influence from each neighbor u surpasses the threshold ϑ_v, the node v will become active, i.e., $\sum_{u \in N_v^a} b_{u,v} \geq \vartheta_v$, where N_v^a is the set of active neighbors of v, and $b_{u,v}$ is the influence weight from u to v.

In independent cascade models, the activation probability is calculated as $p_v = 1 - \prod_{u \in N_v^a}(1 - p_{u,v})$ (Chen et al., 2010; Chen et al., 2009), where $p_{u,v}$ is the influence probability of u on v. It can be seen that the activation probability is calculated based on the assumption that each active neighbor influences v independently. Besides, $p_{u,v}$ is often set as $1/k$ for all $u \in N_v^a$, where k is the total number of active neighbors of v (Chen et al., 2010; Chen et al., 2009), or a constant, such as 0.1 or 0.01 (Kempe et al., 2003; Kimura and Saito, 2006). Such a simple model cannot accommodate the dynamic nature of the adoption process in online social networks, and thus further investigation is needed.

Based on these two classic models, extension has been proposed. In order to capture product information in the social network, Bhagat et al. (2012) propose a linear threshold model with colors, in which both influence of neighbors and product evaluations are incorporated to predict product adoption. Similarly, Chen et al. (2011a) propose a diffusion model based on the independent cascade model, and incorporate negative bias by dividing the active state into two sub-states, i.e., positive and negative.

2.5.2 Influence Maximization

By understanding the diffusion and adoption process, it is important is to predict product adoption. Bhatt et al. (2010) study the adoption of a paid product member of a large and a well-connected instant messenger network for effective online advertising. They find that the spread of adoption remained mostly local to their immediate friends and strong evidence of peer pressure exists in the social networks. Fang et al. (2013) develop a Bayesian learning method to predict adoption probabilities in social networks, in which unobserved confounding factors are included for modeling social influence. Zhang and Pennacchiotti (2013) show that social media profiles from Facebook.com could convey enough information for predicting customers' purchase behavior on eBay.com that can help develop better recommendation engines. In Chapter 7, I propose a linear threshold-hurdle model to describe the product adoption decision making process in a large social network. It overcomes multiple limitations of the current diffusion models, such as modeling influence probability in an ad-hoc fashion, inability to incorporate important operational factors of peer influence, and inability to incorporate the notion of hurdle in the adoption process.

In terms of algorithm development, Kempe et al. (2003) prove that the InfMax problem is NP-hard, and propose a greedy approximation algorithm within $1 - 1/e$ of the optimal influence spread. The performance is significantly better than classic degree and centrality-based heuristics in influence spread. However, this algorithm is still computationally intensive due to the fact that the authors run a sufficient number of Monte-Carlo simulations of the diffusion model to estimate the influence spread. In order to further improve its efficiency, Leskovec et al. (2007) propose a cost-effective lazy forward (CELF) method that exploits the submodularity property of the InfMax problem. Mathematically, function σ is submodular iff $(S \cup \{v\}) - \sigma(S) \geq \sigma(T \cup \{v\}) - \sigma(T)$ whenever $S \subseteq T$, i.e., the marginal gain of influence spread of a new node shrinks as the seed set grows. This property significantly reduces the number of evaluations on the influence spread and thus is 700 times faster than the simple greedy algorithm. Chen et al. (2009) propose fine-tuned heuristics to improve the original greedy algorithm and the CELF method. Experimental results show better efficiency. Further, Chen et al. (2010) design a new heuristics algorithm that is scalable to large social networks (with millions of nodes and edges), controlling the balance between efficiency and InfMax. Goyal et al. (2011) improve CELF by further exploiting submodularity and propose CELF++ which improves the efficiency of CELF by 35%-55%. Goyal et al. (2011) also propose SIMPATH, which improves the efficiency of CELF by incorproating several clever optimizations, including vertex cover optimization and look ahead optimization for InfMax under linear threshold models.

2.6 Marketing-Engineering Coordination

With the thriving of mass customization, many companies aim to expand their product lines and differentiate their product offerings by producing a large number of product variants (Ho and Tang, 1998). From the marketing point of view, this may stimulate sales and increase revenue (Jiao et al., 2007b). This appealing belief does seem true initially. However, when the variety keeps increasing, the engineering consequences of variety explosion emerge, including increasing cost due to complexity increase, inhibiting benefits from economy of scale, exacerbating inventory imbalances, lacking efficiency of manufacturing and distribution (Wortmann et al., 1997). Such engineering concerns must be considered when product design is dominated with profit-based marketing goals.

2.6.1 Product Portfolio Planning

Across industries, product portfolio planning that involves multiple product variants has been the standard practice to reduce cost via economies of scale and scope, reaching multiple market segments and deterring competitors (Simpson, 2004). As mentioned in Chapter 1, it has two main stages, including product portfolio identification and product portfolio evaluation (Jiao and Zhang, 2005a; Li and Azarm, 2002). The former one is to capture and understand customer needs and transform them into product specifications in terms of product attributes and attribute levels. Since I have talked about these topics in Section 2.2, I will mainly review the latter one, i.e., product portfolio evaluation.

One of the most important methods for product portfolio evaluation is the SoC method. It aims to select an optimal level for each product attribute so that the expected number of product adopters can be maximized, for whom the constructed product's holistic utility exceeds a reservation utility or hurdle utility (Camm et al., 2006). Therefore, individual part-worth utilities obtained from conjoint analysis are often used as input to the SoC problem. Due to the fact that different customers have unique preference orderings, the SoC problem involves an optimization problem, that is, the sum of the individual customers' part-worth utilities that indicate the holistic utility of the configured product or product line can maximize its adoption (Wang et al., 2009).

Therefore, two aspects of SoC problem can be identified, i.e., how to obtain individual part-worth utilities with less uncertainty and how to optimize the problem given the individual part-worth utilities. For the first aspect, many researchers use conjoint analysis (Green and Srinivasan, 1978, 1990). Recently, hierarchical Bayesian models based on Markov chain Monte Carlo are used to estimate posterior density of part-worth utilities of individual customers (Allenby and Rossi, 2006; Train, 2003). These methods not only can estimate individual customers' part-worth utilities, but also accommodate uncertainty involved in the estimated part-worth. Therefore, in Chapter 6, I propose to utilize hierarchical Bayesian models based on Markov chain Monte Carlo to estimate part-worth utilities obtained from a CPT-based customer preference model.

On the other hand, the SoC optimization problem is NP-hard, i.e., the problem is unlikely to be solved in polynomial time (Banks et al., 2005). Many algorithms have been proposed. For example, Shi et al. (2012) propose a globally convergent nested partitioning algorithm, incorporating several existing heuristics methods and random sampling form partitions of the feasible regions. Camm et al. (2006) propose an exact algorithm to solve the SoC single-product design problem to provable global optimality, in which logic rules are used to develop and prune the search tree. Later, Wang et al. (2009) propose a branch-and-price approach to the SoC product line design problem and their algorithm is able to identify provably optimal, robust solutions to realistically sized problems. However, in the SoC problem, most of the methods only consider customer preferences to maximize product adoption. Gunnec (2012), on the other hand, incorporates peer influence from the social network in the objective function of the SoC problem, in which peer influence is modeled to decrease a customer's hurdle utility. Differences in market share among the SoC problem with and without peer influence are demonstrated based on their simulation studies. Therefore, it is important to incorporate both customer preferences and peer influence of social networks so as to maximize product adoption in the paradigm of viral product design.

Less attention has been paid to the coordination between engineering and marketing domains in product portfolio planning. Marketers assume that any combinations of product attributes and attribute levels would produce customer preferred products. Such intuitive brief ignores the complex engineering tradeoff that cannot be balanced between customer preferences and engineering constraints (e.g., costs and manufacturability) without coordinating marketing and engineering concerns (Michalek et al., 2011). In this aspect, it is important to consider factors in these two domains at the same time.

2.6.2 Decision-Based Design Approach

Decision-based design is an approach to engineering design that recognizes the substantial role that decisions play in the design and in other engineering activities, largely characterized by ambiguity, uncertainty, risk, and tradeoff (Chen et al., 2013e). Customer preferences in terms of expected holistic utilities are the primary focus in decision-based design (Hazelrigg, 1998). Utility analysis is often used to build mathematical models of a decision maker's preference as a way to identify the optimal option (Thurston, 2006). For example, Orsborn et al. (2009) model preferences for aesthetic forms using a utility function quantitatively. Zhou et al. (2010) model customer preferences in terms of affective needs in accordance with customer emotional satisfaction. These methods aim to maximize customer preferences in terms of a holistic utility measure.

Recently, Chen at al. (2013c) propose an enterprise-driven decision-based design, which attempts to understand the big picture to address enterprise needs as well as attention to the engineering details to meet technical expectations. Under this framework, the typical implementation of decision-based design is profit-driven, involving multiple important questions in marketing, such as cost and price modeling and demand estimation. For example, the cost of a product includes all the possible costs during a product's lifecycle, and thus is very complicated to model, especially in terms of absolute dollars. It is suggested that an estimated distribution of cost is preferred to a point estimate, because a distribution enables inclusion of decision-maker's risk attitudes (Chen et al., 2013b; Thurston and Liu, 1991). The enterprise-driven decision-based design integrates both engineering-level design decisions and marketing-level planning as a single-level optimization, resulting in an all-in-one solution (Du et al., 2014). However, such a single-level optimization strategy is hard to capture the interplay between the domains of engineering design and marketing.

2.6.3 Game Theoretic Formulation

Balancing of marketing and engineering considerations is commonly achieved by integrating the marketing and engineering domains as one singular optimization problem, such that multiple design criteria are aggregated into one "all-in-one" objective function (Luo, 2011). While multi-objective optimization approaches address standalone design problems well, the marketing and engineering interface is related to coupling of multiple decisions, which needs a synergy of conflicting goals of each individual marketing or engineering optimization problem. The all-in-one approach is often practically infeasible in such situations due to computational and organizational complexities (Alyaqout et al., 2011). Optimization by decomposition has been appealing for alleviating the problem of handling a large number of design variables and constraints simultaneously (Kokkolaras et al., 2006). Decomposed optimization largely works only if the domain problem follows a hierarchical decision flow. However, many problems (e.g., product portfolio planning) involving marketing-engineering concerns cannot hierarchically decomposed along disciplinary boundaries. Coordination between marketing and engineering indeed implies equilibrium decisions, whereby different parties strive for different interests and have to compromise with others to achieve common solutions (Devendorf and Lewis, 2011).

In this aspect, one of the potential strategies is bi-level game theoretic optimization. The bi-level optimization problem is originated from the field of game theory with a hierarchical optimization problem (Stackelberg, 1952). It is a special kind of optimization in which a lower-level optimization problem (the follower) is nested within an upper-level optimization problem (the leader) (Colson et al., 2007). In such a structure, the leader and the follower compete against each other so that the leader makes a decision first, and then the follower reacts optimally to the leader's decision as feedback (Du et al., 2014). Based on the follower's decision, the leader adjusts accordingly. The process stops when both obtain satisfactory solutions. As pointed out by Deb and Sinha (2010), the bi-level game theoretic optimization strategy is fundamental to such joint optimization problems, which can deal with the interaction and coupling for conflicting decisions.

However, not many researchers have applied bi-level game theoretic formulations to coordinate marketing and engineering concerns. This is may be because the solution to the bi-level programming is often hard to obtain, mainly due to the non-convexity (Calvete et al., 2008). Even a linear-linear bi-level programming problem is NP-hard (Jeroslow, 1985). Traditional solutions include vertex enumeration, replacing the lower-level problem with its Karush-Kuhn-Tucker conditions (when it is convex and continuous differentiable), and applying gradient methods (Calvete et al., 2008). Nevertheless, these approaches are not technically efficient, especially for large problems, and sometimes lead to a paradox that the follower's decision power dominates the leader's (Lai, 1996). Recently, evolutionary algorithms, such as genetic algorithms (GAs), are promising in dealing with complex optimization problems, as these algorithms usually have a low risk of ending up in a local optimum (Brands and van Berkum, 2014). For example, Calvete et al. (2008) combine classical enumeration techniques with GAs for near-optimal solutions in acceptable computational times. Ji et al. (2013) tackle a leader-follower joint optimization of technical system modularity and material reuse modularity with a constrained GA. Brands and van Berkum (2014) apply a non-dominated sorting GA to solve a bi-level transportation network design problem.

In this research, since I make use of both viral product attributes from the engineering design point view and viral influence attributes from the marketing point view, it needs a coordination process between these two domains. The former (i.e., viral product attributes) implies a product portfolio planning problem, while the latter (i.e., viral influence attributes) implies an InfMax problem. These two problems usually have a certain degree of autonomy and may involve conflicting decisions as well. While multi-objective optimization approaches can well accommodate standalone design problems, they often fail to capture the interaction and coupling involved in the optimization process of product portfolio planning and InfMax. Hence, I formulate viral product design for social network effects as a bi-level game theoretic optimization problem for the purpose of evaluation in Chapter 8. The product adoption maximization problem with viral product attributes and viral influence attributes is modeled as the upper-level optimization problem; the product portfolio planning problem, which aims to maximize customer satisfaction and minimize costs, is modeled as the lower-level optimization problem. Such a bi-level game theoretic optimization is able to capture the interactions and couplings between product portfolio planning and viral marketing.

2.7 Summary

The topics reviewed in this chapter offer guidance to solve the fundamental issues involved in viral product design for social network effects in the next chapter. Considering the limitations of various topics reviewed here, I propose methodologies that can overcome their respective limitations in Chapters 5, 6, 7, and 8 to address a specific step of the viral product design process. For latent customer needs elicitation, I propose use case analogical reasoning from sentiment analysis of online product reviews in Chapter 5. I incorporate subjective experiences, especially affective elements, in customer preference modeling with a descriptive decision choice model, i.e., CPT in Chapter 6. By overcoming multiple limitations involved in the traditional diffusion models, I propose a linear threshold-hurdle model in Chapter 7 for product adoption prediction. Finally, I adopt a bi-level game theoretic formulation (Stackelberg, 1952) that seeks equilibrium solutions between the marketing and engineering domains for viral product design evaluation in Chapter 8.

CHAPTER 3

FUNDAMENTALS OF PRODUCT DESIGN INCORPORATING PEER INFLUENCE OF SOCIAL NETWORKS

Recognizing the importance of social network effects in product design, this chapter examines the fundamental issues underlying product design incorporating peer influence of social networks, from the research domains reviewed in the previous chapter, including engineering design, viral marketing, social network analysis, and marketing-engineering coordination. Understanding these fundamental issues is crucial to this research and thus paves the way for a new paradigm of design in the next chapter.

3.1 A Holistic View

The main question of this research is how to incorporate peer influence of social networks into product design. From different topics in different research areas reviewed in Chapter 2, I summarize the fundamental issues in terms of viral attributes, customer preferences modeling, diffusion mechanism, adoption maximization, and marketing-engineering coordination. Starting from the interactions among customers, products, and social networks, their interrelationships are illustrated in Figure 3.1. Viral attributes include both viral product attributes and viral influence attributes. The choice set of viral product attributes are identified by latent customer needs elicitation, based on which customer preferences can be modeled. Viral influence attributes make use of the social network effects, resulting in peer influence of social networks. Based on a diffusion model, I can predict product adoption, considering customer preferences, peer influence, and hurdle utility. Finally, by coordinating marketing-engineering concerns, it aims to maximize product adoption. Viral product attributes contribute to both customer preferences and product virality, which form the sufficient cause for product adoption. However, the necessary cause is the influence of social network effects, without which a product would not be diffused and adopted by a large number of customers. How to leverage these two important aspects and marketing-engineering coordination is crucial to the success of product design incorporating peer influence of social networks.

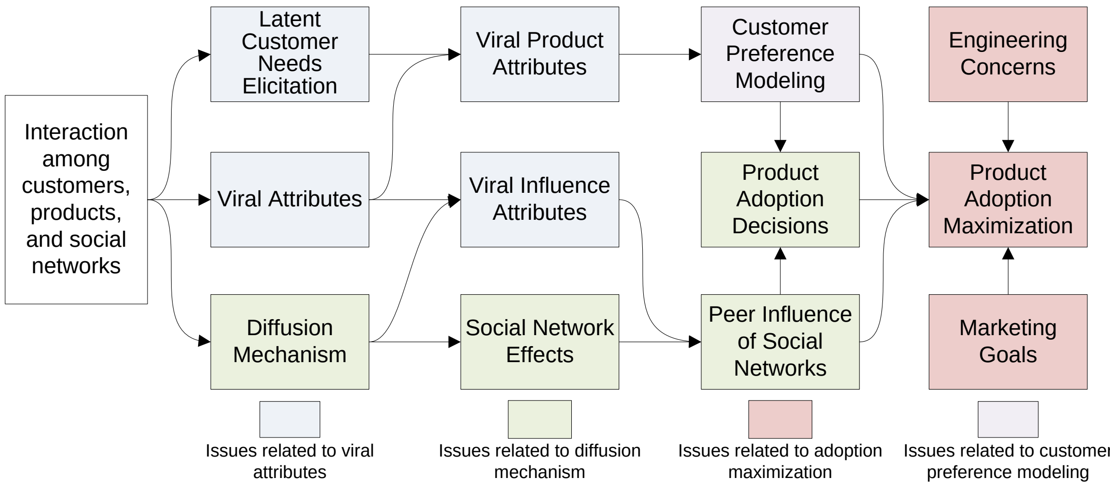


Figure 3.1 Fundamental issues involved in product design incorporating peer influence

3.2 Viral Attributes

We distinguish two types of viral attributes: viral product attributes A^{VP} and viral influence attributes A^{VI}. Thus, viral attributes are the union set of them, i.e., $A^{V} = A^{VP} \cup A^{VI}$.

3.2.1 Viral Product Attributes

Among the set of product attributes A, there are certain attributes more prone to social network effects, and thus are defined as viral product attributes, i.e., $A^{VP} \subseteq A$. Viral product attributes are fundamentally about the content of a product and the psychological effects that can have on a customer's desire to share the product with peers (Stephen and Berger, 2009). A product's viral influence attributes, on the other hand, concern how the product is shared - how such attributes catalyze a product's awareness and adoption in relation to other customers in the social network. Research has shown that psychological characteristics of content play an important role in shaping virality. In their study, Berger and Milkman (2010) find that positive, useful, and surprising news from New York Times tend to be viral than others. One good example is probably the vanity mirror mounted on the sun visor in the car. When it was first introduced, it was among the most popular dealer-added accessories that provided high profit margins with the sales (Genat, 2004). It became viral among women and then among all the passengers. However, as time goes by, it becomes a basic product attribute in the car. This is consistent with the Kano model (Kano et al., 1984) in that product attributes which can satisfy latent customer needs tend to be viral, because these product attributes lead to unexpected delight and satisfaction. However, it gradually evolves into a basic product attribute over time.

3.2.2 Viral Influence Attributes

Viral influence attributes are embedded in the social network rather than in the product. They may enable communication, generate automated notifications of customers' activities, facilitate personalized invitations, or enable hypertext embedding of the product on publicly available websites and weblogs. Two types of the most widely used viral influence attributes are personalized referrals and automated broadcast notifications (Aral and Walker, 2011). For example, in their study, Aral and Walker (2011) find that passive-broadcast viral messaging generates a 246% increase in local peer influence, while adding active-active-personalized viral messaging only generates an additional 98% increase in contagion, despite the fact that active-personalized messaging is more effective in encouraging adoption per message.

Another type of viral influence attributes is seed customers that correspond directly to the viral mechanism associated with the viral marketing aspect of the product. Seed customers offered incentives by firms are able to influence the largest possible number of other customers to adopt a product. These seed customers are thus influential and are often the opinion leaders or social hubs within a social network (Goldenberg et al., 2009). For example, celebrities in Twitter (https://twitter.com/) tend to be seed customers as they usually have a huge number of followers who can be possibly influenced by their celebrities. Thus, how to identify seed customers is important to the influence maximization (InfMax) problem. Other possible viral influence attributes include tagging to prompt to share, commenting on like or dislike, and inputting customer reviews or satisfaction levels, and so on.

3.3 Customer Preference Modeling

The aim of customer preference modeling is to quantify customer preferences. Usually it is achieved by a holistic product utility U_{ij} indicates customer v_i's preference to the product P_j. A holistic product utility is often represented as a weighted sum of individual part-worth utilities. By making direct utility assessment over the product attributes, it assumes that part-worth utility functions are independent of each other (Zhou and Jiao, 2013b). However, customers' interactions with a product (characterized of multiple product attributes) result in a holistic impression of perceived value on every individual product attribute (i.e., part-worth utilities). Such a belief of customer preferences can hardly be consistent with the mindset of composing a weighted sum of single assessments of individual product attributes (Zhou et al., 2010). This suggests that part-worth utilities actually are dependent on one another, due to the coupling of customer interactions with multiple product attributes. Therefore, the integration of multiple product attributes necessitates the incorporation of an uncertain dependence analysis as an integral part for design decision support.

Moreover, individual part-worth utilities are often obtained by conjoint analysis without considering any subjective experience and risk attitudes in choice decision making as shown in Chapter 2 (Zhou et al., 2014b). Customers often compare one product with another before they make choices so that the decision is often reference-dependent, which conforms to the human perceptual process (Tversky and Kahneman, 1992). In the context of social networks, different people may have very different attitudes and subjective experiences towards even the same type of peer influence. While some may be prone to be influenced by his social network peers, others may be indifferent. Such dynamics and irrationality coincides with risk attitudes of decision making. Therefore, it is important to improve the utility preference measure by formulating prospect value functions that excel in accounting for risk attitudes and subjective experiences of decision makers (Zhou et al., 2014b).

3.4 Diffusion Mechanism

The dynamics in social networks makes the diffusion mechanism complicated. Rogers (2003) summarize important findings about diffusion of innovations. Diffusion is the process in which an innovation (e.g., Kindle Fire HD tablets) is spread through certain channels over time among social entities in a social system. Four important factors are identified to influence the spread, including the innovation itself, the diffusion channels, time and a social system. Usually the innovation must be widely adopted in order to self-sustain. This is related to the threshold theory (Granovetter, 1978; Granovetter and Soong, 1988) in social science. With regard to adopters, five different categories are recognized, including innovators, early adopters, early majority, late majority, and laggards. Other factors such as cultures, customs and rituals also influence diffusion of innovations. In this research, important issues related to diffusion mechanism in the context of online social networks are identified as follows:

3.4.1 Peer Influence

We have identified three effects in the social networks that are contributing to information diffusion in Chapter 1, including word-of-mouth effects, imitation effects, and network effects (Dou et al., 2013). We aggregate the influence of these effects and understand them as peer influence. Aral (2011) defines peer influence in the social network from the utility point of view as “how the behaviors of ones’ peers change the utility one expects to receive from engaging in a certain behavior and thus the likelihood (or extent to which) that one will engage in that behavior.” First, the sources of peer influence are the peer behaviors perceived by the social entity. These behaviors influence the social entity to adopt a product by changing their utility function about the product. Second, Aral contends that peer influence is causal, which excludes correlated and confounding effects. However, I argue that in this research influence of social network effects are broader than peer influence, which can include correlational effects, such as herding behavior and homophily (McPherson et al., 2001), i.e., the tendency of individuals to link to others in the social networks. This is because similar choice behavior may result from similar opinions, interests, and attitudes towards products and services (Aral et al., 2009). Therefore, it is useful to predict product adoption based on the similarity between two social entities.

3.4.2 Threshold Theory

Another important theory related to diffusion mechanism is the threshold theory proposed by Granovetter and his colleagues (1978; 1988). It states that individuals’ behavior depends on the number of other individuals already engaged in that behavior (termed as a behavioral threshold). Similar to the definition of peer influence by Aral, the individuals often use a utility function to calculate his/her cost and benefit from undertaking an action based on the current situation. In other words, individuals' thresholds relate to the utility of joining the collective behavior or not. The threshold model has been used to explain riots, residential segregation, and the spiral of silence, as well as other applications, including social contagion (Valente, 1996). Therefore, in the context of online social networks, it needs a threshold or a user base before the spread of a product can take off. With the increasing use of Internet and especially social networks, the user base is often reached within a relatively short time by paid word-of-mouth seed customers for instance. As long as the user base has reached, by capitalizing on the social networks they reside in, words of mouth can dramatically increase product awareness or achieve other marketing objectives, such as product sales, through self-replicating viral processes, similar to the spread of virus or computer virus (Howard, 2005). Due to the fact that online social networks are no longer location restricted, online social network users are provided with greater exposure and visibility to product information than ever before (Darren et al., 2012).

3.4.3 Operational Factors

In order to understand social network effects for product adoption maximization, one important question is how to identify the operational factors underlying social network effects. Besides the viral product attributes, I mainly focus on social entities and the social network structure to understand social network effects. Rice (1990) observe that social interactions among people represent an important force influencing individual's adoption behaviors in a social network. Social influence network theory (Friedkin, 1998) also posits that a person endowed with an initial opinion or behavioral assessment receive and respond to information propagated in a social network and could choose to modify an original opinion or assessment accordingly. By studying the structure of social networks, Granovetter (1973) argues that weak ties (people loosely connected to others in the network) are necessary for diffusion to occur across subgroups within a system. Burt (1987) presents a third network approach to diffusion by arguing that structural equivalence (the degree of equality in network positions) influences the adoption of innovations. Bampo et al. (2008) also examine the impact of the social structure of social networks (random, scale free, and small world) and of the transmission behavior of individuals on viral marketing campaign performance. Their simulation studies show that scale-free networks are very efficient for viral campaigns, and thus it is possible for campaign managers to capture scale-free properties in their target audience by identifying and seeding influential customers who might then function as hubs. Other social network measures, which may influence product diffusion, include centrality, density, and reciprocity (Rice, 1994).

3.4.4 Activation Threshold

The activation threshold is derived from the threshold theory (Granovetter, 1978; Granovetter and Soong, 1988) that if the aggregated peer influence is larger than a customer's activation threshold, he or she will become active. Due to individual differences, the activation threshold is often modeled as a uniform distribution between 0 and 1 in diffusion models (Kempe et al., 2003). This is not consistent with the studies in the domain of innovation diffusion and communication. As mentioned previously, Rogers (2003) identifies five categories of adopters in diffusion research, including innovators, early adopters, early majority, late majority, and laggards. Innovators are the very first individuals to adopt a product, who are willing to take risks and thus have the lowest activation thresholds. The second earliest category of individuals who adopt the product is early adopters, who often have a higher social status and advanced education, and are more socially forward than late adopters. Hence, they also have relatively lower activation thresholds compared with late adopters. Early majority tends to be slower in the adoption process, and seldom holds positions of opinion leaders in a social system. Thus, they tend to have higher activation thresholds than innovators and early adopters, but have lower activation thresholds than late majority and laggards. Late majority often has a high degree of skepticism about the product, and has relatively high activation thresholds. Laggards typically have an aversion to change, and thus have the highest activation thresholds in the social system. Apparently, each category of adopters does not have an activation threshold distributed uniformly between 0 and 1.

3.4.5 Hurdle Utility

One limitation in traditional diffusion models is that once a customer is activated by peer influence, he or she will adopt the product unconditionally (e.g., Chen et al., 2010). As matter of fact, activation is not tantamount to adoption in the context of viral product design (Bhagat et al., 2012). For example, a friend of mine from Facebook experienced the new iPhone 6 Plus from the Apple store and thinks it is so great and powerful. I indeed agree with that, but think it is too expensive. Then I share my opinion on Facebook. Some of my friends get influenced, and do not buy the product. From this example, I am activated by my friend on Facebook without buying iPhone 6 Plus, and still influence my other friends in the social network. From this example, activation is similar to product awareness. Kalish (1985) characterizes the process of product adoption in terms of two steps, i.e., awareness and adoption, and argue that the awareness of information, i.e., product diffusion, spreads in an epidemic-like fashion, whereas the actual adoption depending on other factors. Given customers' needs, preferences, and social network effects, they act so as to make a satisficing decision of whether to adopt the product or not (Simon, 1956). Different individual customers have different hurdle utilities (Camm et al., 2006; Wang et al., 2009), which essentially perform as a reservation price – the highest price an individual is willing to pay for (the holistic utility of) a product. Therefore, product diffusion is often regarded as a proxy for product adoption (Bhagat et al., 2012), and if the joint utility of product preferences and peer influence is larger than the hurdle utility, the customer is expected to adopt the product.

3.5 Adoption Maximization

In order to maximize product adoption by incorporating peer influence of social network in the product design process, it is important to understand two related problem, namely the share-of-choice (SoC) problem in product portfolio planning and InfMax in viral marketing.

3.5.1 The Share-of-Choice Problem

Product portfolio planning (i.e., product line design) aims to select a near-optimal mix of product variants configured by different product attribute levels to offer in the target market (Jiao et al., 2007b). A product line involves multiple product variants (e.g., 16GB, 32GB, and 64GB models of Kindle Fire HD tablets), which attract heterogeneous customers in the market. Each product variant is configured with K product attributes, i.e., $a_k \in A, k = 1, \dots, K$. Let L_k be the number of levels for the k-th attribute. Then the attribute level set can be represented as $A^* = \{a_{kl}^* | k = 1, \dots, K, l = 1, \dots, L_k\}$. Considering all the possible configurations, there are a number of meaningful J product configurations, indicated by $X = \{X_j | j = 1, \dots, J\}$, and $X_j = (X_{jk} | k = 1, \dots, K) = (x_{j11}, \dots, x_{j1L_1}, \dots, x_{jK1}, \dots, x_{jKL_K})$ is a vector showing a particular configuration for the *j*-th product P_j, where $x_{jkl} | k = 1, \dots, K, l = 1, \dots, L_k$ is 1 if the *l*-th attribute level of the *k*-th attribute is selected for P_j, otherwise it is 0. $Y = (y_j | j = 1, \dots, J)$ is a vector indicating a particular choice of P_j, and $y_j = 1$ if P_j is chosen, otherwise it is 0.

In the literature of product portfolio planning, the basic principle of identifying an optimal set of product attributes and attribute levels is the SoC problem, aiming to maximize the number of customers who will adopt the product (Kohli and Krishnamurti, 1989). For a set of customers, $V = \{v_1, v_2, \dots, v_N\}$, they have different part-worth utilities, i.e., u_{ikl}, which denotes the utility of the *l*-th attribute level of *k*-th attribute perceived by the *i*-th customer in the social network. The SoC adoption model postulates that a customer v_i adopts the *j*-th product only if his/her holistic utility of the product U_{ij} exceeds his hurdle utility, h_{ij}. In this situation, $z_{ij} = 1$, otherwise it is 0. The objective of the SoC problem is to maximize the number of product adopters, i.e., $\sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}$, denoted as $\sigma(X)$. Therefore, the SoC problem can be formulated as follows (Du et al., 2014; Gunnec, 2012):

$$\text{Max} \quad \sigma(X) = \sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}, \tag{3.1a}$$

$$\text{s.t.} \quad U_{ij} = \sum_{k=1}^{K} \sum_{l=1}^{L_k} u_{ikl}\, x_{jkl} \geq h_{ij} z_{ij}, \forall i \in [1, N], j \in [1, J], \tag{3.1b}$$

$$\sum_{l=1}^{L_k} x_{jkl} = 1\,, k \in [1, K], j \in [1, J], \tag{3.1c}$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L_k} \left| x_{jkl} - x_{j'kl} \right| > 0\,, j \neq j', \tag{3.1d}$$

$$\sum_{j=1}^{J} y_j \leq J^+ < J, \tag{3.1e}$$

$$z_i = \sum_{j=1}^{J} z_{ij}\,, j \in [1, J], \tag{3.1f}$$

$$z_{ij}, \in \{0,1\}, x_{jkl} \in \{0,1\}, \forall i \in [1, N], j \in [1, J], k \in [1, K], l \in [1, L_k]. \tag{3.1g}$$

The exclusive constraint (3.1c) guarantees that only one attribute level is selected for each product attribute. Constraint (3.1d) denotes that at least one attribute level is different for two different product variants. J^+ in constraint (3.1e) is the upper bound for the total number of product variants offered in the market.

3.5.2 Influence Maximization

The InfMax problem in viral marketing aims to select a set of key customers to act as the seeds to influence the largest number of customers so that they are likely to become product adopters in a social network (Kempe et al., 2003). Given a directed social network $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_N\}$ is a set of nodes, indicating a set of customers and a link from v_j to v_i, i.e., $(v_j, v_i) \in E$ means that v_j can potentially influence v_i. Given a certain budget, the key is to find a seed set $S \subseteq V$ with $|S| = n$, so that by activating them, I can maximize the expected number (denoted as $\sigma(S)$) of customers that eventually get activated based on a diffusion model. Note these seed customers are considered as one kind of viral influence attributes, as they are often offered incentives to promote a product by ways including positive review, personal referral, and broadcasting. Assume I_i^t is the influence of social network effects of v_i from his or her active neighbors at time t. I_i^t can be calculated by aggregating individual influence from active neighbor v_j to v_i, i.e., I_{ji}^t at time t. In order to activate v_i, I_i^t has to be no smaller than his or her activation threshold θ_i. When v_i is activated, it leads to product adoption, i.e., $z_{ij} = 1$ if the product is P_j, otherwise it is 0. Then the adoption problem is to find S^* that maximizes $\sigma(S)$ (Zhou and Guo, 2014),

$$\text{Max} \quad \sigma(S) = \sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}, \tag{3.2a}$$

$$\text{s.t.} \quad I_i^t = \sum_{\substack{(v_j, v_i) \in E \\ v_j \text{ is active}}} I_{ji}^t \geq \theta_i \quad, \forall i \in [1, N]\,, \tag{3.2b}$$

$$|S| = n, \tag{3.2c}$$

$$S \subseteq V. \tag{3.2d}$$

3.5.3 Adoption Maximization with Viral Attributes

From the formulation of the SoC problem and the InfMax problem, they share the same objective to maximize the number of product adopters (see Eq. (3.1a) and Eq. (3.2a)) and entail a threshold-based constraint (see Eq. (3.1b) and Eq. (3.2b)). However, the SoC problem considers a set of customers V only, without taking into account their relationships (i.e., the social network), whereas the InfMax problem considers all of the products to be the same (Barbieri and Bonchi, 2014). The underlying assumption of the InfMax problem ignores the fact that customers may have different preferences to different product attributes. Moreover, influence spread is not equivalent to adoption spread (Bhagat et al., 2012). Other factors that influence actual adoption, such as price and individual's valuation of the product, are not captured in classic diffusion models (Lu and Lakshmanan, 2012). This can be indicated as the hurdle in the SoC problem. Therefore, I propose to integrate these two decision problems in terms of product design incorporating peer influence of social networks by introducing the adoption probability (Barbieri and Bonchi, 2014):

$$\Pr\left(P_j \middle| i,t\right) = \frac{\exp\left(\varphi_i^t(P_j)\right)}{1+\exp\left(\varphi_i^t(P_j)\right)} \geq \vartheta_i, \tag{3.3}$$

where

$$\varphi_i^t\left(P_j\right) = I_i^t + U_{ij} - h_{ij}, \tag{3.4}$$

where $\Pr\left(P_j \middle| i,t\right)$ is the probability that v_i will adopt product P_j at time $t+1$ if it is no smaller than the adoption threshold ϑ_i at time t. This formulation enables both viral attributes and social network effects to be incorporated in product adoption modeling, and thus referred to as adoption maximization with viral attributes (AdpMaxVA).

$$\text{Max } \sigma(S, X), \tag{3.5a}$$

$$\text{s.t.} \quad \Pr\left(P_j \middle| i,t\right) \geq \vartheta_i, \forall i \in [1,N], j \in [1,J], \tag{3.5b}$$

$$S \subseteq V, |S| = n, X \subseteq A. \tag{3.5c}$$

3.6 Marketing-Engineering Coordination

The interplay of the InfMax problem and the SoC problem necessitates a joint optimization, i.e., AdpMaxVA. It leverages a set of viral influence attributes (i.e., seed customers) and a set of viral product attributes, and thus need to coordinate both marketing and engineering concerns. In the marketing domain, particularly viral marketing, the goal is to maximize profits or market share by maximizing product adoption. In the engineering domain, product designers aim to reduce the cost of their products while offering a set of product variants that can be preferred by heterogeneous customers. Therefore, only these two objectives are well-coordinated, an enterprise-level goal can be possibly obtained. Unlike the all-in-one formulation in decision-based design (Chen et al., 2013e) as introduced in Chapter 2, it is important to capture the interaction and interplay between these two domains.

In the context of social networks, it is important to leverage social network effects as a medium to coordinate between the two domains. As shown in Figure 2.2, the output of engineering design according to its own objectives is different product variants (i.e., a product line), which are subject to peer influence of social networks. The main contributing factors in terms of a product are the viral product attributes. Another important factor is the viral influence attributes, such as a set of seed customers. These factors with the influence of social network effects lead to the resulting product adoption. Compared with the viral marketing goal and constraints, a new set of design requirements are generated to revise the product line, based on which a new product line will be marketed in the social network to test the marketing goal. Such a coordination process is expected to reach an equilibrium solution between the engineering domain and the marketing domain.

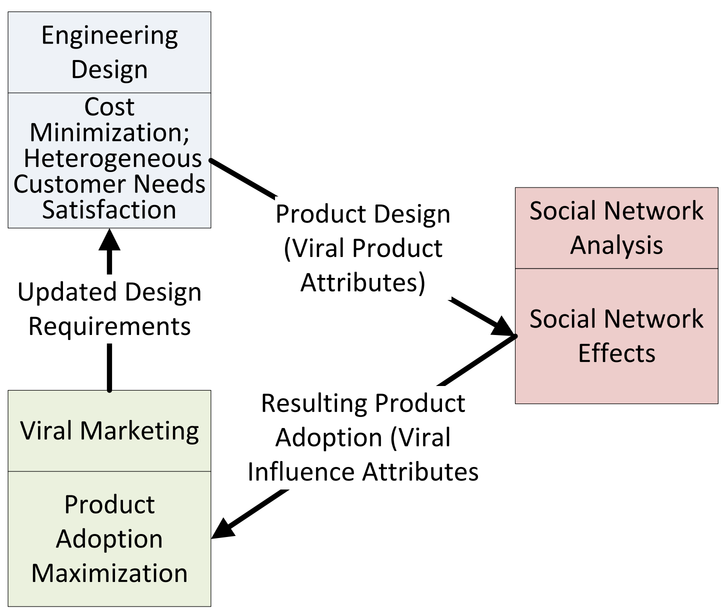


Figure 3.2 The coordination process between engineering and marketing

3.7 Product Design Incorporating Peer Influence

Based the fundamental issues identified above, I formulate the problem of product design incorporating peer influence as follows:

Given a social network $G = (V, E)$ and social network data S^D,

---

Find an optimal set of J^+ optimal product configurations and an optimal set of n seed customers in the social network that maximize product adoption and optimize product line performance,

Subject to 1) a set of identified product attributes and attribute levels obtained from S^D;

2) Customer preferences with regard to a particular product;

3) The diffusion process in the social network;

4) Marketing-engineering coordination.

3.8 Summary

This chapter examines the fundamental issues underlying product design incorporating peer influence of social networks. These fundamental issues include identification of viral attributes, customer preference modeling, the diffusion mechanism of a social network, and adoption maximization, as well as engineering-marketing coordination. Their interrelationships are also elaborated, and a formulation of product design incorporating peer influence of social networks is also presented. Such a profound understanding of these fundamental issues provides us a clear direction for a technical approach in the next chapter.

CHAPTER 4

VIRAL PRODUCT DESIGN FOR SOCIAL NETWORK EFFECTS

In order to deal with the fundamental issues of product design incorporating peer influence of social networks, viral product design for social network effects is proposed with a technical research framework, which forms the theoretic foundation of this research. A coherent four-step design process is presented along the technical research framework, including latent customer needs elicitation for identifying candidate viral product attributes, customer preference modeling and quantification accommodating subjective experiences, social network modeling for product adoption prediction, and bi-level game theoretic optimization for viral product design evaluation. The technical challenges of these four design steps are identified and the corresponding methodologies are subsequently proposed.

4.1 A New Paradigm of Design

To answer the fundamental issues of product design incorporating peer influence, there may be multiple approaches and methodologies. Among many, I propose a new paradigm of design, i.e., *viral product design for social network effects*. It aims to provide specific design guidelines for product adoption maximization in the context of online social networks, considering marketing-engineering coordination. First of all, social network effects are the cause of peer influence in the social network. How to leverage and capitalize on peer influence is the main focus in this paradigm of design. Second, viral attributes, including both viral product attributes and viral influence attributes can be designed in such a way that product adoption can be maximized and product line performance can be optimized at the same time under the constraint of marketing-engineering coordination. Third, this paradigm of design is situated in the context of online social networks. Therefore, I assume that the input is the comments and reviews of a product (line) within an online social network, whereas the output is a set of viral product attributes and a set of seed customers as viral influence attributes. The former with other possible product attributes configures a set of product variants, i.e., a product line, which satisfies the objective of this paradigm of design, i.e., maximizing product adoption, while satisfying engineering concerns, such as maximizing customer satisfaction and minimizing costs. The latter is a set of influential seed customers, and by offering incentives to them, such as discounts, coupons, and other promotion strategies, they are able to influence a maximum number of customers to adopt the product.

4.2 A Technical Framework for Viral Product Design

To drive the development of rigorous research methodologies of viral product design for social network effects, a technical framework is presented in Figure 4.1 using IDEF0 (Icam DEFinition for Function Modeling) (Grover and Kettinger, 2000). In essence, the technical framework integrates three major domains, including engineering design, marketing, and social computing, and account for four coherent design steps, namely,

A0.1) Latent customer needs elicitation for viral product attributes extraction,

A0.2) Customer preference modeling and quantification for product choice decision making,

A0.3) Social network modeling for product adoption prediction, and

A0.4) Viral product design evaluation by adoption maximization.

As shown in A0, the input is online product reviews and comments and social networks, and the output is 1) the optimal product configurations formed by selected viral product attributes and 2) identified influential seeds from the social network. Influential seeds are considered as one kind of viral influence attributes that can make use of peer influence of social networks to influence other non-adopters to become adopters. Both the viral product attributes and influential seeds are expected to maximize product adoption and optimize product line performance jointly. On the upper side, control factors, including customer needs, customer characteristics, subjective experiences, and social network effects, are used to constrain the design steps. On the lower side, resources are used to complete the task, including designers, customers, and company stakeholders. Among them, customers adopt products, review and comment products, and form social networks; designers identify viral product attributes and influential seeds based on the goals made by company stakeholders.

A0 is further decomposed into four steps. The first step (A0.1) is to elicit latent customer needs and extract customer desired product attributes (and attribute levels) with sentiment analysis from online product reviews and comments. Note I emphasize latent customer needs so that the product attributes that satisfy these needs are more likely to be viral in the social networks. These desired product attributes form the choice set of viral product design.

The product attributes and attribute levels are the input for customer preference modeling and quantification in the second step (A0.2), based on which their individual part-worth utilities can be obtained. In addition, I incorporate subjective experiences, including affective states, cognitive tendency, and risk attitudes in the choice decision making process for customer preference modeling. The output is the perceived product prospect, indicating the holistic product utility.

The third step (A0.3) is social network modeling for product adoption prediction. Its input is the holistic product utility and the social network. As two important customer characteristics, i.e., activation threshold and hurdle utility are used to constrain the prediction results, namely, adoption or not.

The fourth step (A0.4) is to maximize product adoption and optimize product line performance jointly for viral product design evaluation. Its input is the social network and product attributes produced from the first step. The output from A0.3, i.e., the adoption result, becomes the control factor for the bi-level decision making process. The resources are the influential seed customers in the social network that are used to influence other customers so that the largest possible number of product adopters can be expected. The output are viral product attributes (or optimal product configurations) and influential seeds that are identified.

These attributes configure product variants in the product line that are supposed to be viral in the social network. Customers will further review and comment on the products and their tastes may change as time goes by, especially for technological products, such as smart phones. The influential seeds identified residing in the social network may also change due to the evolution of the social network. Thus, the whole process form an iterative design process that is ought to improve the design and maximize product adoption continuously.

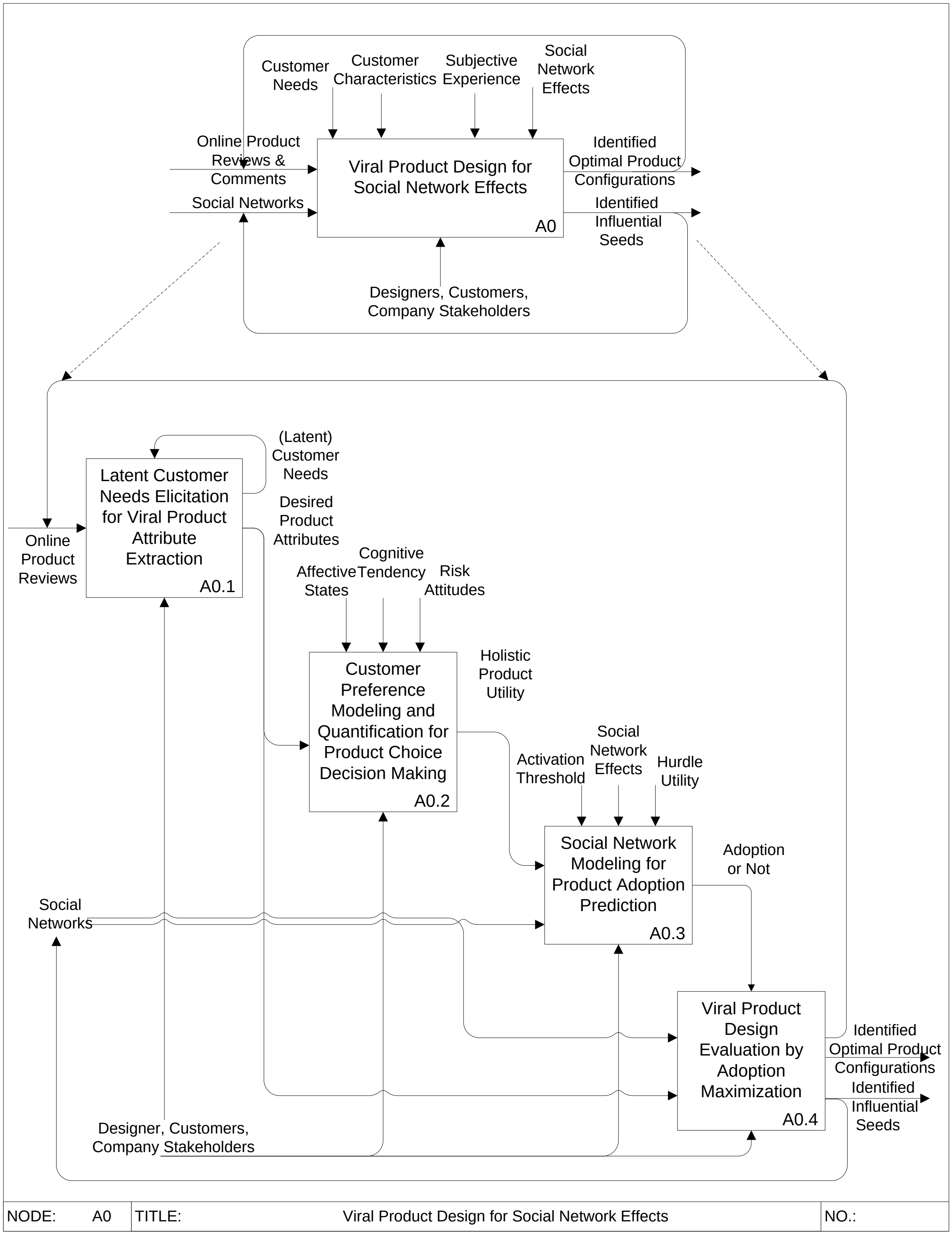


Figure 4.1 A technical framework of viral product design for social network effects

4.3 Technical Challenges

4.3.1 Latent Customer Needs Elicitation

In order to identify product attributes that may be viral, it is important to identify latent customer needs. According to the Kano model (Kano et al., 1984), there are three kinds of product attributes, including must-be quality, one-dimensional quality, and attractive quality. Among them, attractive quality attributes provide unexpected delight for customers. These attributes corresponding to latent customer needs, which may be non-obvious and very difficult to identify (Otto and Wood, 2001). Sometimes, customers may not be consciously aware of them, but are surprised or delighted if they are satisfied (Wagner and Hansen, 2004). Traditionally, empathic design methods for examining the customer in the natural environment have great potential in identifying latent needs, such as customer observation (Hanski et al., 2014). However, the observation data may be biased based on the interpretation of design engineers and analyzing the data is time consuming. Another important method is to involve customers in the design process actively (Zhou et al., 2013). Despite many advantages and innovations of customer co-design, it may also be time-consuming and costly to identify, recruit, and understand the right customers and their roles.

Another challenge for latent customer needs is their linguistic analysis. Customer needs are often expressed in linguistic terms, which are often abstract, fuzzy, or conceptual (Tseng and Jiao, 1998). Although much effort has been devoted to leveraging tools and technologies (e.g., text mining) to enhance customer needs elicitation, still there is much space for improvement (Meth et al., 2013). Furthermore, it is often difficult to create a high-quality information channel that runs directly among the customers in the target market, the researchers in marketing, and the designers of the product. People in different domains often express the needs in different set of contexts and differences in semantics and terminology always impair the ability to convey needs information effectively from customers to marketing folks and to designers (Jiao and Chen, 2006).

4.3.2 Customer Preference Modeling and Quantification

Traditional customer preference and choice models, such as conjoint analysis and discrete choice analysis, do not explicitly incorporate subjective experiences, especially affective elements. In the domain of decision sciences, prevailing computational models for analyzing and simulating human perception and evaluation on choice decision making are mainly cognition-based models (Ahn, 2010). However, recent affective neuroscience and psychological studies have reported that human affect and emotional experience play a significant role in human decision making (Ahn and Picard, 2005; Bracha and Brown, 2012). Furthermore, the interaction between the customer's cognitive aspect and affective aspect often makes it complicated in the decision making process (Zhou et al., 2011c). Another main challenge to understand customer choice decision making is to design appropriate measures with both construct validity and predictive power with regard to the evaluation of a product (Law and Van Schaik, 2010). Traditional methods use the concept of part-worth utilities for attribute levels. The question is how to aggregate the part-worth utilities so that the holistic product utility can be well represented.

4.3.3 Social Network Modeling

In order to model social networks for product adoption prediction, it is important to understand how the dynamics of adoption are likely to unfold within the underlying social network. As described in Chapter 2, two types of diffusion models (Kempe et al., 2003), i.e., independent cascade models and linear threshold models, are proposed. One problem is that the adoption probability is often modeled in an ad-hoc way. Assuming v is an inactive social entity (i.e., non-adopter) at step t, and the probability of v becoming active (i.e., adopters) depends on the influence of v's neighbors who are adopters at step t. In the linear threshold model, each node v has an activation threshold θ_v uniformly distributed between 0 and 1. When the joint influence by each neighbor u surpasses the threshold ϑ_v, the node v will become active, i.e., $\sum_{u \in N_v^a} b_{u,v} \geq \theta_v$, where N_v^a is the set of active neighbors of v, and $b_{u,v}$ is the influence weight from u to v. In independent cascade models, the activation probability is calculated as $p_v = 1 - \prod_{u \in N_v^a}(1 - p_{u,v})$ (Chen et al., 2010; Chen et al., 2009), where $p_{u,v}$ is the influence probability of u on v. It can be seen that the adoption probability is calculated based on the assumption that each active neighbor influences v independently. Besides, $p_{u,v}$ is often set as $1/k$ for all $u \in N_v^a$, where k is the total number of active neighbors of v (Chen et al., 2010; Chen et al., 2009), or a constant, such as 0.1 or 0.01 (Kempe et al., 2003; Kimura and Saito, 2006). Such a simple model cannot accommodate the dynamic nature of the adoption process in online social networks, and thus further investigation is needed.

Furthermore, when a social entity adopts a product, it is assumed that he or she will influence others to adopt it as well in a positive way in the classic diffusion models (Bhagat et al., 2012). Obviously, this is not entirely true. A certain percentage of customers will give negative reviews about the product, if it cannot satisfy their needs. Therefore, it is important to incorporate negative product reviews that discourage other users to adopt the product in the diffusion process. Another assumption that may not hold is that only adopters can share their user experience about a product and influence their neighbors subsequently (Bhagat et al., 2012). In some shopping websites, I observe that many people give comments about certain products without purchasing the product, such as Amazon.com. These non-adopters may be influenced by other adopters, and act as an information bridge in the product diffusion to influence others in the social network.

4.3.4 Viral Product Design Evaluation

In order to evaluate viral product design, it is important to understand and solve the optimization problem involved in viral product design. The interplay of the SoC problem and the InfMax problem necessitates a joint optimization problem, that is, to optimize product line performance and maximize product adoption that leverages both marketing and engineering concerns. The maximum expected number of product adopters and optimal product line performance thus become the measures for viral product evaluation. Traditional approaches often integrate the marketing and engineering domains as one single optimization problem, such that multiple design criteria are aggregated into an “all-in-one” objective function (Luo, 2011). While multi-objective optimization approaches address standalone design problems well, the marketing and engineering interface is related to coupling of multiple decisions, which needs a synergy of conflicting goals of each individual marketing or engineering optimization problem. The all-in-one approach is often practically infeasible in such situations due to computational and organizational complexities (Alyaqout et al., 2011). Coordination between marketing and engineering indeed implies equilibrium decisions, whereby different parties strive for different interests and have to compromise with others to achieve common solutions (Devendorf and Lewis, 2011).

4.4 Technical Approach

Corresponding to the technical framework and the technical challenges discussed above, I propose the overall technical approach as shown in Figure 4.2.

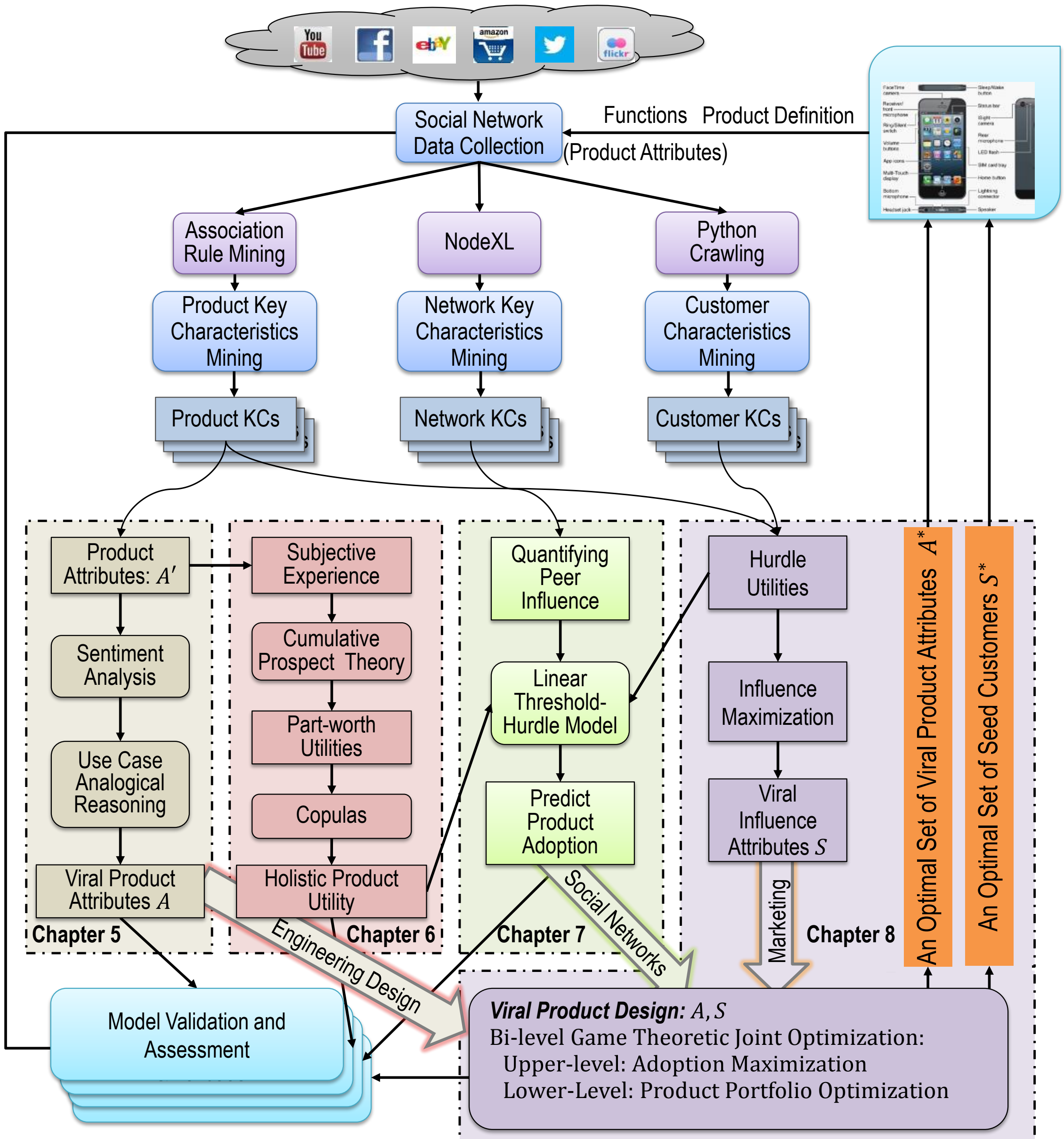


Figure 4.2 Technical approach—viral product design for social network effects

The overall solution strategy consists of three parts, including data collection, methodologies, and model validation and assessment. Data collection is to investigate the potential of social network mining for analyzing key characteristics of the social networks in terms of the product, network, and customer key characteristics. Note that the inputs to social network search and data collection are textual terms used to describe various functions of the product under study. Also social network data collection is conducted by searching for the terms of individual functions, instead of the product itself. For example, when I want to design a new iPhone X, it is not available in the market yet, and definitely there is little information about this new product in social media for us to understand product adoption decisions. If I break it down and search for the textual terms related to its functions in terms of product features, such as “camera”, “touch screen”, and “battery life”, I would find plenty of useful information in social media. Indeed, this is consistent with the typical market and requirement analysis task in the product definition stage of design.

Different methodologies to solve the research issues involved in the four steps identified in Section 4.3 are shown in Figure 4.2, which are corresponding to Chapters 5 to 8. Viral product attributes are extracted using use case analogical reasoning from sentiment analysis of product key characteristics; Holistic product utility is obtained by aggregating individual part-worth utility based on cumulative prospect theory (CPT). A linear threshold-hurdle model is proposed to predict product adoption based on the network key characteristics. Viral product design evaluation is formulated as a bi-level game theoretic joint optimization problem. The results are two optimal sets of viral product attributes and seed customers.

Finally, corresponding to each methodology, a case study will be conducted for the purpose of model validation and assessment.

4.4.1 Use Case Analogical Reasoning from Sentiment Analysis

Given the challenges mentioned above, I propose to elicit latent customer needs by use case analogical reasoning from sentiment analysis of online product reviews as described in Chapter 5. The rise of social media, such as blogs, social networks, and review websites (e.g., Amazon.com and Eopinion.com) has fueled the proliferation of reviews, ratings, recommendations, and other forms of online expressions.

(1) Sentiment analysis: It is the computational study of opinions, sentiments, and emotions expressed in online texts (Liu, 2010), and three important steps are needed, including 1) What product attributes are evaluated, i.e., attribute extraction, 2) What is the evaluation polarity/orientation, i.e., sentiment prediction, and 3) Aggregating sentiments over each product attribute. For example, I do not like the charger of my tablet, which expresses a negative opinion on the charger of the tablet. Of all the customer reviews, the number of positive opinions and the number of negative opinions indicate the trend of customer preferences about this attribute. By analyzing the negative opinions about the charger, I can understand customer needs. For example, if the charging process is too long, then the corresponding customer need is to shorten the charging process. For online product reviews, the majority are expressed with emotional tones, making it easy to identify customer likes and dislikes. Despite the fact that the quality of the review data varies, many reviews have detailed opinionated information about specific product attributes in terms of their quality, usability, and aesthetics, etc. Therefore, sentiment analysis can be conducted with regard to specific product attributes rather than the general product (e.g., Chen et al., 2012; Hu and Liu, 2004b; Jin et al., 2009). Such links between the opinionated information and product attributes give direct clues for eliciting customer needs.

(2) Use case analogical reasoning: Another kind of important information that sentiment analysis can extract is different use cases from online product reviews. Compared with ordinary use cases, extraordinary ones tend to create a context to elicit latent customer needs. This is consistent with the key idea in (Lin and Seepersad, 2007), in which extraordinary use cases are created to break the mold of users' thought process and usage pattern, thereby encouraging them to interact with the product in innovative ways and to articulate latent needs that lead to breakthrough products. In this sense, the users interacting with the product in extraordinary use cases become empathic lead users (Hannukainen and Hölttä-Otto, 2006). Similarly, Chen et al. (2013a) propose usage context-based design, in which product performance varies significantly under different use cases. This greatly affects customer preferences and choices. However, I am not able to interview these users. What I can obtain is their online product reviews. We propose to apply case-based reasoning (CBR) (Goel and Craw, 2005; Zhou et al., 2011a; Hayes et al., 2011) for use case analogical reasoning to elicit latent needs. CBR is the process of solving new problems based on the solutions of similar past problems, and is one of the most powerful methods for analogy reasoning (Riesbeck and Schank, 1989; Aamodt and Plaza, 1994). In this research, customer needs are elicited for ordinary cases that are mined through sentiment analysis of online product reviews. Whenever an extraordinary case is identified, the most similar case in the database will be retrieved. The elicited customer needs for the retrieved case are reused to elicit latent needs for extraordinary use cases by case adaptation.

4.4.2 Prospect Theoretic Modeling of Customer Preference

In this research, I 1) propose a customer preference-based decision making model based on CPT, 2) incorporate the influence of affect in the model as an extension to the original CPT theory, 3) estimate parameters involved in the model as a way to test the hypotheses and accommodate both individual heterogeneity and group homogeneity, and 4) aggregate individual part-worth utilities with copulas that considers interdependence among them (Zhou et al., 2014b; Zhou and Jiao, 2013b).

(1) Prospect theoretic modeling of customer preference: I adopt CPT as the basic customer preference-based decision making model. Prospect theory (Kahneman and Tversky, 1979) is originally proposed to describe decision making in the domain of behavioral economics. Unlike normative decision models (e.g., expected utility theory), CPT is a descriptive decision model that describes how humans actually make decisions as it is rather than it ‘should’. A CPT value function is defined with respect to a reference point, rather than in terms of an absolute value as in expected utility theory, and thus is reference dependent. Such an emphasis on the reference point conforms to the human perceptual process, which tends to notice shifts more than resting on static states (Tversky and Kahneman, 1992). Furthermore, CPT modifies the original prospect theory by applying the probability distortions to the cumulative probabilities so that stochastic dominance is not violated. Customer preference choice decision making essentially shares the same process with those in behavioral economics. Thus, CPT provides a "legitimate" method to evaluate customer preferences to product configurations.

(2) Affective influence: We incorporate the influence of affect in the decision making process by shaping the parameters involved in the CPT-based preference model. The original CPT model does not incorporate the influence of affect. However, according to Ahn (2010), the parameters of the value function change systematically in sequential decision-making situations, involving incidental affective states and task-related confidence. Therefore, it is possible to incorporate affective factors by shaping the CPT parameters in customer preference modeling, whereby human choice behavior is made regarding multiple design attributes; the attribute level corresponding to neutral or indifference can be regarded as a reference point; Those above and below the reference point can be formulated to express positive customer preferences and negative customer preferences, respectively; and affective influences and cognitive tendencies can be incorporated in the value function and in the weighting function by parameter shaping.

(3) Parameter estimation: In order to estimate multiple parameters associated with the model, a hierarchical Bayesian parameter estimation procedure is developed, taking affective influence into account. The prevailing method to estimate model parameters is either to average data across groups of participants to uncover underlying patterns or to use individual data to accommodate individual differences. However, the former cannot accommodate individual differences and the latter usually has few and thus noisier data, based on which the result is more unreliable compared to the former (Harrison and Rutström, 2009). In this research, a hierarchical Bayesian model is proposed to estimate the parameters involved in the customer preference model. It offers a principled and comprehensive way to relate psychological models to experimental and observational data (Lee and Newell, 2011). It can identify how the variables are related, inferring causal influences between customer preferences and product attributes that go beyond regression or correlation analysis (Steyvers et al., 2009). Meanwhile, a hierarchy of sub-models is useful to analyze the data with one model at the group level for studying customer homogeneity and another model at the individual level for studying customer heterogeneity.

(4) Aggregating part-worth utilities with copulas: In order to model the dependence among the product attributes, I propose an extension of utility copulas, i.e., nested utility copulas, to formulate multivariate utility copulas for aggregation of dependent part-worth utilities (Zhou and Jiao, 2013b). Nested utility copulas are based on the concept of modular design, in which different product attributes are first grouped into different modules so that the dependence between product attributes of within-modules will be high and those of between-modules will be low. This formulation further simplifies the construction of multi-attribute utility functions with multivariate utility copulas.

4.4.3 Linear Threshold-Hurdle Model

In order to address the limitations discussed in Section 4.3.3, I propose a linear threshold-hurdle (LTH) model to describe product adoption decision making process in the context of online social networks. Based on the LTH model, a data mining method, i.e., rough set is used to predict product adoption (Zhou et al., 2014c).

(1) LTH model: First, I model the activation threshold based on the five categories of adopters, and each has different uniform distributions. Second, three operational factors are identified to model peer influence probabilities within the social networks, including interaction strength, entity similarity, and structural equivalence. Third, if a social entity's influence probability is larger than its activation threshold, it will become active. The social entity will adopt a product on condition that his/her perceived holistic product utility surpasses his/her hurdle utility, otherwise he or she will enter a state called tattle, in which either a positive or negative review is expressed. Similarly, after the social entity adopts the product, he or she will also comment the product in a positive or negative way. These positive or negative reviews and comments can further influence other social entities' perceived holistic product utilities.

(2) Rough set prediction model: Then, a data mining method named rough set (Pawlak, 1991) is used to predict whether a social entity will adopt a product or not based on the features extracted from the LTH model. The rough set theory has been widely applied in the domain of data mining for the purpose of prediction (Zhou et al., 2014a; Zhou et al., 2011b). It excels at tackling vagueness and uncertainty using rough approximations by producing a complete set of consistent and minimal decision rules, using an objective knowledge induction process (Pawlak, 1991).

4.4.4 Bi-Level Game Theoretic Optimization

In order to deal with viral product design evaluation, I adopt a bi-level game theoretic approach and formulate the marketing and engineering coordination corresponding to these two optimization problems as a Stackelberg game (Stackelberg, 1952). It seeks for an equilibrium solution between maximizing product adoption and optimizing product line performance jointly in the marketing and engineering domains, respectively.

(1) Joint optimization for product adoption and product line performance: I propose to construct a leader-follower joint optimization model to reveal the bi-level game theoretic decision structure underlying marketing-engineering coordination in viral product design. It essentially entails a bi-level stochastic programming (Colson et al., 2005; Luo et al., 1996), consisting of an upper-level optimization problem (leader) and a lower-level optimization problems (follower). In this research, the leader F leverages both viral product attributes and influential seeds for product adoption maximization as the marketing goal. The basic idea is to identify an optimal set of product attributes and attribute levels and an optimal set of seed customers that can maximize product adoption in the social network. The follower f leverages product portfolio planning. It considers both customer satisfaction and engineering cost that can maximize the expected overall performance of the entire product line as the engineering goal. Between the leader and the follower, the social network is where the decision is tested. A leader-follower model assumes certain decision power for both the leader and the follower, with the leader possessing a higher priority.

(2) Hybrid Taguchi-genetic algorithm: Being generally non-convex and non-differentiable, bi-level programs are intrinsically hard to solve. Even the simplest instance, like the linear-linear bi-level programs, has been shown to be NP-hard (Vicente and Calamai, 1994). Existing solutions mostly involve a lower level that admits extreme solutions, a property that allows the development of methods that guarantee a global optimum. When the lower-level problem is convex and regular, it can be replaced by its

Karush-Kuhn-Tucker conditions, yielding the single-level reformulation with a Lagrangian function associated with the lower-level problem. While the Lagrangian constraint is linear in certain cases (linear or convex quadratic functions), the complementarity constraint is intrinsically combinatorial, and can be addressed by enumeration algorithms, such as branch-and-bound (Camm et al., 2006; Thoai et al., 2002), or branch-and-pricing (Wang et al., 2009).

Due to a large variety of attribute configurations and seed options, enumeration algorithms tend to be difficult to be computationally tractable. Among many, genetic algorithms (GAs) are promising over other methods to solve bi-level programs, such as dynamic programming, beam search, greedy heuristics, or simulated annealing (Belloni et al., 2008; Gunnec, 2012). To alleviate complexity of enumeration, my strategy is to take advantage of a generic structure inherent in product differentiation (Jiao and Tseng, 1999). To better cater to the context of product portfolio planning and social network analysis, I will explore the potential of developing efficient, flexible, and robust GAs.

Specifically, a hybrid Taguchi-genetic algorithm (HTGA) is proposed. It is based on orthogonal GAs, in which some major steps (e.g., crossover) of a GA algorithm can be considered as "experiments" and the Taguchi's method or orthogonal design can be used to identify the global optimum more robust and statistically sound (Zhang and Leung, 1999). The Taguchi's method (Taguchi, 1995) can produce orthogonal arrays, i.e., fractional factorial design, which can scan the feasible space more evenly to locate good points for further exploration in the following iterations (Leung and Wang, 2001). Tsai et al. (2004) further introduce the notion of signal-to-noise ratio (SNR) to select better offspring, which further increases the quality and robustness of a GA, termed as the HTGA. In this research, I follow their idea and extend the original 2-level orthogonal arrays to n-level ($n > 2$) ones in the context of a bi-level optimization problem.

Since the bi-level optimization problem involves two types of variables, viral product attributes and seed customers. I adopt a coordinate-wise optimization strategy (Barbieri and Bonchi, 2014) to fix one type of variables (i.e., seeds or product attributes) in the optimization problem. When the set of product attributes is fixed, it becomes an InfMax problem. I adopt an improved greedy algorithm (Goyal et al., 2011) to solve the InfMax problem. When the seed set is fixed, I propose a HTGA method to solve the bi-level optimization problem. First, I transform the bi-level problem into a single-level parametric optimization problem. This transformation guides us to the solution with the HTGA method, in which the fitness function can be obtained with a penalty function. Second, the HTGA method makes use of the Taguchi method in experiment design to improve the robustness and efficiency in the chromosome generation process. Thus, the interaction and coupling of these two optimization problems are addressed with the coordinate-wise optimization strategy iteration by iteration, and the optimal set of product attributes and seed customers are identified for viral product design.

4.5 Summary

This chapter proposes viral product design for social network effects as a technical approach to product design incorporating peer influence of social networks. Corresponding to the fundamental issues identified in the previous chapter, I propose a technical framework with four coherent design steps, including A0.1) Latent customer needs elicitation for viral product attributes extraction, A0.2) Customer preference modeling and quantification for product choice decision making, A0.3) Social network modeling for product adoption prediction, and A0.4) Viral product design evaluation by adoption maximization. Along with the technical framework, the main challenges and possible solutions for each design step are identified. The detailed problem formulation, methodology, and experimental and case studies related to these design steps are presented in Chapters 5, 6, 7, and 8, respectively.

CHAPTER 5

LATENT CUSTOMER NEEDS ELICITATION BY USE CASE ANALOGICAL REASONING FROM SENTIMENT ANALYSIS

Customer needs elicitation is an integral part of product design. Satisfying customers' needs, especially latent customer needs, can delight customers unexpectedly, which leads to product adoption to a large extent. Furthermore, product attributes extracted from the customer elicitation process become the choice set of viral product attributes. However, traditional methods often confront many technical challenges, including time-consuming and costly data collection, ambiguous linguistic analysis of customer needs, and inability to identify latent customer needs, and so on.

This chapter proposes a new method of latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews. First, fuzzy support vector machines (SVMs) are used to build sentiment prediction models with different kernel functions. These models are further built on the features extracted from a list of affective lexicons based on affective norms for English words and WordNet. Second, product attributes and use cases are mined with association rule mining and are refined by term similarity measures. Thus, sentiment analysis is able to summarize sentiment and comment frequency information on individual product attributes. Third, inspired by analogical reasoning, I capitalize on case-based reasoning (CBR) to reuse and adapt ordinary use cases so as to elicit latent customer needs for extraordinary use cases. A case study of Kindle Fire HD 7 inch tablet is used to illustrate the potential and feasibility of the proposed method.

5.1 Latent Customer Needs for Viral Product Attributes Extraction

The success of a product or a service is largely dependent on to what extent the product or the service satisfies customer needs, including solving customers' problems and making them feel good. Many companies strive to offer customer-focused products and services with a large degree of individuality for the purpose of customization and personalization (Zhou et al., 2013). Poor understanding of customer needs and inaccurate assumptions made during the elicitation and analysis of customer needs have a significantly negative impact on the design and manufacturing of the product in terms of quality, lead time, and cost (Jiao and Chen, 2006). Therefore, products that cannot satisfy customer needs lead to poor product adoption. Negative words of mouth can spread in the social network, which makes the product hard to survive in the social network.

Generally speaking, the steps of eliciting customer needs consist of 1) gathering raw data from customers, 2) interpreting the raw data into customer needs, 3) organizing the needs into a hierarchy of primary, secondary, and (if necessary) tertiary needs, 4) prioritizing the needs with relative importance, and 5) reflecting on the results and the process (Ulrich and Eppinger, 2003). Usually multiple methods can be used to gather raw data, such as interviews, focus groups, observations, processing tracing methods (e.g., think-aloud), and other conceptual methods (e.g., laddering) (Crandall et al., 2006). Different methods often involve different amounts of time with varied effectiveness. For example, it is reported that 90% of the customer needs for picnic coolers are revealed after 30 interviews and 98% of customer needs for a piece of office equipment are identified after 25 hours of data collection with both interviews and focus groups (Ulrich and Eppinger, 2003). When translating the raw data into customer needs, usually customer needs are expressed in terms of product attributes with enough details. However, not all of the translation can be done directly from the raw data without further interpretation. Correspondingly, the general needs become primary needs, and each of which will be characterized as a set of secondary needs and even tertiary needs in the same fashion. This process is still very laborious if no design support tools are used. In order to prioritize customer needs, many techniques can be applied to assign importance weights to them, such as quality function deployment and conjoint analysis (Prasad et al., 2010), which often need to collect more data from customers to set up experiments. The final step is to verify that the customer needs collected are consistent with the knowledge and intuition of the development team (Ulrich and Eppinger, 2003).

Compared with explicit customer needs, latent customer needs are much harder to elicit. This is due to the fact that latent needs may be non-obvious and very difficult to identify (Otto and Wood, 2001), and sometimes customers may not be consciously aware of them. Nevertheless, customers are surprised or delighted if they are satisfied (Wagner and Hansen, 2004). One important method to elicit latent customer needs is to interview lead users who experience needs still unknown to the public (Hannukainen and Hölttä-Otto, 2006). For example, by putting users in situations so that they are hard to see or hear, the use case makes users interact with products in extraordinary ways (Hannukainen and Hölttä-Otto, 2006). In this sense, those who experience needs unknown to the public in the extraordinary cases are regarded as lead users (Von Hippel, 1986) in this research. Accordingly, these needs are considered to be latent with regard to the majority of the public. Specifically, in this chapter,

(1) I propose sentiment analysis of online product reviews that greatly facilitates the processes of data collection and linguistic analysis in customer needs elicitation through text mining. A combination of affective lexicons and fuzzy SVMs is used to predict customer preferences with regard to individual product attributes from online user-generated product reviews.

(2) I propose association rule mining to extract major product attributes and attribute levels with importance measures as well as use cases, which are further refined with term similarity measures to reduce redundancy.

(3) I capitalize on use case analogical reasoning with CBR to elicit latent needs from extraordinary use cases by reusing and adapting ordinary use cases. Extraordinary use cases from online product reviews are extracted from sentiment analysis. By extraordinary I mean the use case is transformed into an interaction situation, in which customer needs are not known to the majority of the public. Thus these needs are considered to be latent and are elicited by customizing and adapting analogical customer needs from ordinary use cases.

5.2 Problem Formulation

Figure 5.1 shows the steps of latent customer needs elicitation from online product reviews, including data collection, attribute/case identification, sentiment prediction, and use case analogical reasoning.

(1) The first step is data collection by a data crawler tool, Python 2.7 (www.python.org), in which the input is a set of web pages from a review website (e.g., Amazon.com, Epinions.com), $R^P = \{r_i^p\}_{i=1}^{I}$, where I is the total number of the web

pages, and the output is a file of product reviews, $F_r = \{r_i\}_{i=1}^{R}$, where R is the total number of the reviews.

(2) In the second step, association rule mining is first used to extract a set of product attributes, $A' = \{a'_k\}_{k=1}^{K'}$ that customers reviewed, and then a similarity matching algorithm is used to generate a set of refined product attributes, $A = \{a_k\}_{k=1}^{K}$. K' and K are the total numbers of the original product attributes and refined product attributes, respectively, and $K' \leq K$. Each attribute can assume a number of levels, i.e., $A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$, where L_k is the total number of levels of a_k. This same process is also conducted to identify refined cases (see Section 5.7.3), i.e., $C^R = \{C_k^R\}_{k=1}^{M_R}$, where M_R is the total number of refined cases.

(3) In the third step, a model based on fuzzy SVMs is used to predict sentiments of product reviews with regard to individual product attributes, based on which customer preference information can be obtained, i.e., $u_k = \{a_k, s_k, f_k\}_{k=1}^{K}$, where s_k is a sentiment variable, which can have values of positive sentiment or negative sentiment, $s_k = \{p^s, n^s\}$, and f_k is a variable indicating comment frequency among all the reviews. Corresponding to the attribute levels, their respective customer preferences can also be expressed as $u_{kl} = \{a_{kl}^*, s_{kl}, f_{kl}\}_{l=1}^{L_k}$.

(4) In the last step, CBR is used for analogical reasoning between ordinary and extraordinary uses cases in order to elicit latent customer needs. Design engineers are involved to scrutinize elicited latent customer needs for the purpose of consistency with respect to each design attribute, i.e., $C_k^N = \left\{a_k, \{c_{kj}^n\}_{j=1}^{J_k}\right\}_{k=1}^{K}$, where J_k is the total number of customer needs of a_k. Note all the steps are guided by the stakeholders and the company goals. Therefore, the problem of latent customer needs elicitation is formulated:

Given a set of I review web pages $R^P = \left\{r_i^p\right\}_{i=1}^{I}$,

---

How to identify product attributes $A = \{a_k\}_{k=1}^{K}$, attribute levels,$A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$, and use cases $C^R = \{C_k^R\}_{k=1}^{M_R}$,

How to predict sentiments of product reviews for customer preference elicitation $u_{kl} = \{a_{kl}^*, s_{kl}, f_{kl}\}_{l=1}^{L_k}$,

How to elicit latent customer needs $C_k^N = \left\{a_k, \{c_{kj}^n\}_{j=1}^{J_k}\right\}_{k=1}^{K}$,

Subject to constraints of stakeholders and company goals.

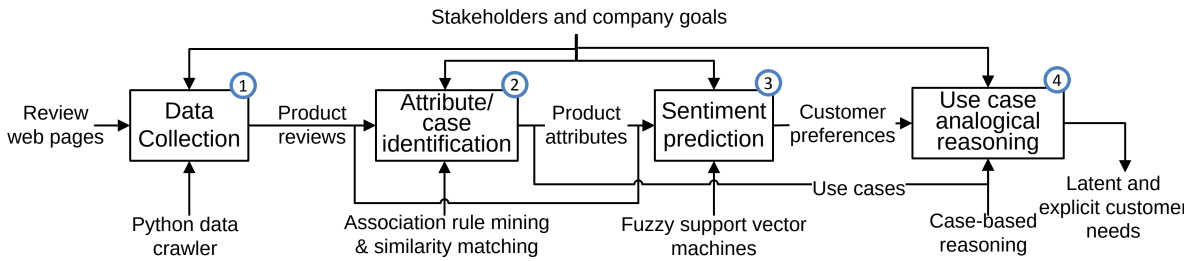


Figure 5.1 Steps involved in latent customer needs elicitation

5.3 System Architecture

Corresponding to the problem formulation in Figure 5.1, Figure 5.2 shows the system architecture for use case analogical reasoning from sentiment analysis of online product reviews with five modules, namely data pre-processing, lexicon construction, attribute/case extraction and refinement, fuzzy SVM model building & testing, and use case analogical reasoning for latent customer needs elicitation.

(1) Data pre-processing: The purpose of this module is to process raw data for the subsequent sentiment analysis and latent customer needs elicitation. Python is used to crawl online reviews and removes meaningless symbols (e.g., HTML symbols). The cleaned reviews are segmented into sentences and these sentences are labeled with either positive or negative with the help of user-provided ratings.

(2) Lexicon construction: This module is to construct a list of affective lexicons. ANEW (Affective Norms for English Words) (Bradley and Lang, 1999) is first used as lexicon seeds with quantitative scores of valence, arousal, and dominance, and then WordNet (Miller, 1995) is used to expand the sentiment lexicon seeds based on a standard label propagation algorithm (Blair-Goldensohn et al., 2008).

(3) Attribute and case extraction & refinement: In this module, product attributes commented by the reviewers are first extracted based on association rule mining (Hu and Liu, 2004b). Then a similarity measure is used to cluster product attributes that are described with similar terms (e.g., photo and picture taken by a camera) by a matching algorithm. Different levels of attributes are then categorized together in a hierarchical structure with the help of design engineers. Meanwhile, this module also extracts different use cases (both ordinary and extraordinary cases; see Figure 5.8) with association rule mining. The refinement process is similar to that of attribute refinement.

(4) Fuzzy SVM model building & testing: This module is to train a fuzzy SVM model for sentiment prediction and validation. First, based on the lexicons and labels, the fuzzy SVM model with different kernel functions are built and the parameters in SVMs are tuned with a cross-validation strategy. Based on these parameters, the model is tested for sentiment prediction with a 10-fold cross-validation process.

(5) Use case analogical reasoning for latent customer needs elicitation: Based on the product attributes extracted and their sentiment information, customer opinions are summarized, facilitating the process of translating customer opinions into explicit customer needs for ordinary use cases. For the extraordinary use cases, ordinary use cases are reused and customized to elicit latent customer needs with CBR for analogical reasoning. Based on the review statistics, I prioritize design attributes as a way to assign importance weights to customer needs. Finally, an evaluation and reflection process on the results is conducted.

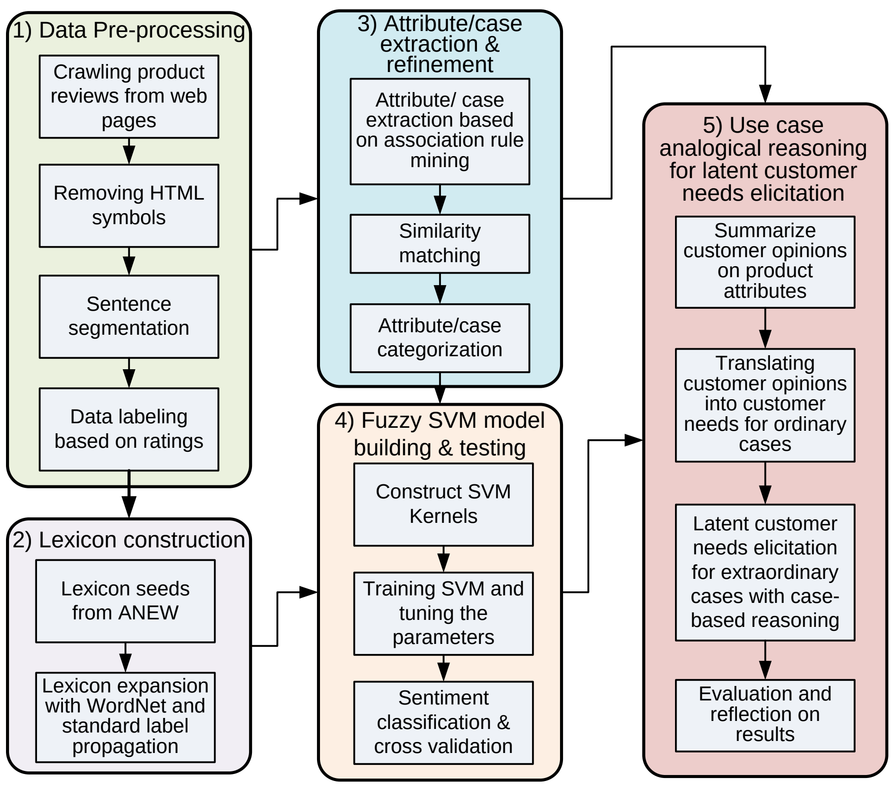


Figure 5.2 The system architecture of latent customer needs elicitation based on use case analogical reasoning from sentiment analysis

5.4 Case Study

In order to illustrate the proposed method, a Kindle Fire HD 7 inch tablet (first generation, released on September 14, 2012) is used as a case study. The review data are collected from Amazon.com using Python 2.7 from October 2, 2012 to November 20, 2013 with user-provided ratings (from 1 star to 5 stars). The total number of review pages $I = 1973$, and each page has 10 reviews, which result in $R = 19730$ reviews. Figure 5.3 shows a typical review about this product, including its helpfulness votes, rating (5 stars), the time, the name of the reviewers, the location, the product he/she commented, the review content, and the number of comments on the review.

8 of 11 people found the following review helpful
love my kindle but..., October 6, 2013
By **donna l browning** - See all my reviews
This review is from: Kindle Fire HD 7", HD Display, Wi-Fi, 16 GB - Includes Special Offers Electronics)
the kindle fire hd ss awesome. i had wanted one for the lingest ti.e and finally got one. i already had a prime account because of my granddaughter`s kindle. I`ve read a lot of free books and have downloaded many free apps. The only shortcoming I`ve noticed is the battery life. For instance, i unplugged it from the charger arou d 8 this morning add tt was 100% I just picked it up after not using it all day and battery is at 38% so I`m pretty much charging it every day. That ss the only negative. I love everything else about it. My husband has an ipad and i much prefer the kindle over the ipad.
Help other customers find the most helpful reviews Report abuse | Permalink
Was this review helpful to you? Yes No Comments (2)

Figure 5.3 A typical review of Kindle Fire HD 7 inch tablet

These reviews are then segmented into sentences. After attribute extraction, the sentences that do not contain product attributes are discarded. If a product attribute is described in multiple sentences, only one sentence is kept for the purpose of normalization. For those sentences that contain product attributes, I extract the data features (see Section 5.5.2), based on which SVM models are used to predict polarity of the sentence that describes the product attribute(s). The ground truth is obtained by the customer ratings provided, i.e., those with ratings of 1 star or 2 stars are considered as negative and those with ratings of 4 stars or 5 stars are considered as positive. Those with ratings of 3 stars account for 20.6% of all the reviews and are discarded. This is because these reviews tend to have mixed opinions for the product under evaluation, which often complicates the labeling process. A total number of 4475 opinionated sentences (3210 positive sentences and 1265 negative sentences) that contain product attributes are generated from 500 randomly selected reviews that are longer than 5 sentences.

5.5 Sentiment Prediction

5.5.1 Lexicon Construction and Propagation

A general sentiment lexicon is built on ANEW and WordNet. The words (1033 in total) in ANEW are rated against three dimensions between 1 and 9, including valence, arousal, and dominance. For example, the word 'arrogant' is rated as 3.69 (2.40), 5.65 (2.23), and 5.14 (2.71) for three respective dimensions for all the subjects (Note 3.69, 5.65, and 5.14 are mean values while 2.40, 2.23, and 2.71 are standard deviations). Please refer to (Bradley and Lang, 1999) for more information. Of the three dimensions, the mean rating values of each word of all the subjects are used in this study. Thus, all the 1033 words are regarded as seed lexicons of known ratings on three dimensions, and are then expanded through synonym and antonym links in WordNet (Miller, 1995). One limitation of ANEW is that no part-of-speech tags are available. In order to overcome it, the expanded list in WordNet is accompanied with simple part-of-speech tags, including noun (n), verb (v), adjective (a), and adverb (r). This is done to avoid ambiguities raised by polysemy. For example, bar.v (as a verb) is a synonym of barricade.v or banish.v while the noun bar.n (as a noun) is not. Unlike the method in (Hu and Liu, 2004a), not only a list of sentiment words is created, but also each word is measured against valence, arousal, and dominance that represent how likely the given word has the designated positive or negative sentiment. The construction process of sentiment lexicons is described below:

First, the 1033 words from ANEW are used to find their synonyms and antonyms from WordNet using Python, resulting in a total number of 10713 words. Second, the expanded list of sentiment lexicons weighs against valence, arousal, and dominance, respectively. The algorithm is modified from the standard label propagation (Blair-Goldensohn et al., 2008; Zhu and Ghahramani, 2002). For the k-th word in ANEW, w_k $(k = 1, \ldots, 1033)$, I identify its synonym set and antonym set, denoted as syn(w_k) and ant(w_k), respectively. Assuming that there are s elements in syn(w_k) and a elements in ant(w_k), they form a set about w_k, indicated as $s(w_k) = \{w_k^j\}_{j=1}^{n_w}$, where n_w is the total number of elements in set $s(w_k)$ with the first element as itself (i.e., w_k). Then, I initialize a score vector s_k^0 with $n_w = 1 + s + a$ elements as $s_k^0 = [x_k, 0, \dots ,0]^T$, where x_k is the corresponding value of valence, arousal, or dominance of w_k, and the remaining is a series of $s + a$ zeroes. Note that the values of valence, arousal, and dominance are normalized between -1 and 1. Define

$$s_k^{m_w} = A_p^{m_w} * s_k^0, \tag{5.1}$$

where the propagation matrix $A_p = [a_{ij}]_{n_w \times n_w}$, and $a_{ij} = \begin{cases} 1 + \lambda, \text{if } i = j \\ \lambda, \text{if } w_k^i \in \text{syn}\left(w_k^j\right) \\ -\lambda, \text{if } w_k^i \in \text{ant}(w_k^j) \\ 0, \text{otherwise} \end{cases}$. Here I set m_w = 5 and $\lambda_w = 0.2$, since larger values of m_w do not improve performance and larger values of λ_w lead to too skewed a distribution of the scores, according to (Blair-Goldensohn et al., 2008). Finally, $s_k^{m_w}$ is also normalized as follows:

$$s_{k\_n}^{m_w} = s_k^{m_w} / \text{abs}(\max(s_k^{m_w})) * \text{abs}(x_k), \tag{5.2}$$

where $\text{abs}(\cdot)$ is a function to obtain its absolute value. An example of the 750th word 'radiant' in ANEW is used to illustrate the algorithm described above. Assume that its synonym set and antonym set are {'beaming.a', 'beamy.a', 'bright.a'} and {'dull.a'}, respectively, and beaming.a and beamy.a are synonymous with each other. Then $s(w_{750})$ = {radiant.a, beamy.a, beaming.a, bright.a, dull.a}, and $s_{750}^0 = [0.496, 0, 0, 0, 0]^T$, where 0.496 is the normalized valence value of 'radiant'. Note I assign a part-of-speech tag to 'radiant' based on the part-of-speech tags of its synonyms and antonyms. The final score is $s_{750\_5}^5 = [0.496, 0.299, 0.299, 0.224, -0.224]^T$ after 5 iterations and normalization. Likewise, I can obtain their normalized values of arousal and dominance. Of all the scores I obtained, most of them are consistent with affective semantics if not all, and Table 5.1 gives more examples derived from the proposed method.

Table 5.1 Examples of lexicons with part-of-speech tags and scores of valence, arousal, and dominance

Words	Valence	Arousal	Dominance
easily.r	0.5529	-0.0043	0.2658
endow.v	0.5175	0.2488	0.1322
promotion.n	0.8052	0.4222	0.2491
powerless.a	-0.5163	-0.2593	-0.2678
recession.n	-0.0309	-0.1306	-0.0420
superb.a	0.6431	0.2014	0.2068
suspicious.a	-0.1371	0.3242	-0.0028

5.5.2 Prediction Based on Fuzzy Support Vector Machines

(1) Data feature extraction: Since the lexicon list cannot cover all the English words, a backup opinion lexicon list is provided with 2006 positive words and 4783 negative words (Hu and Liu, 2004a). Simple valence values are assigned to them (i.e., 0.5 to all the positive words and -0.5 to all the negative words). Besides, it includes useful properties for online texts, such as misspellings, morphological variants, slang, and social media markups. In this research, I choose 10 features from a sentence by a sequential feature selection method among 14 proposed features, described as follows:

1) Number of words with positive valence, i.e., N_{pos};

2) Number of words with negative valence, i.e., N_{neg};

3) The average of valence, arousal, and dominance, i.e., $V_{ave}, A_{ave}, D_{ave}$;

4) The maximum valence and corresponding arousal and dominance, i.e., V_{max}, A, D, if $N_{pos} > N_{neg}$; The minimum valence and corresponding arousal and dominance, i.e., V_{min}, A, D ,if $N_{pos} \leq N_{neg}$;

5) Number of negation words, i.e., N_n;

6) Number of words that denote adversative relations, i.e., N_a;

The features in 1) and 2) are extracted by string comparison between the target sentence and the affective lexicon lists with Python, while the features in 3) and 4) are obtained from ANEW and the lexicon propagation algorithm described in Section 5.5.1. The features in 5) and 6) are also obtained by string comparison between the target sentence and a predefined list of negation words (e.g., no, never, none) and a predefined list of adversative words (e.g., but, nevertheless), respectively. The last feature is useful when the sentence has a mixed sentiment. For example, "the screen of the tablet is wonderful, but the battery charger sucks." In such a case, I will further divide the sentence into two sub-sentences, i.e., "the screen of the tablet is wonderful" and "the battery charger sucks." This helps accurately predict customer preferences in terms of individual product attributes.

(2) Model construction and prediction: We apply fuzzy SVMs (Lin and Wang, 2002) with different kinds of kernel functions to predict sentiment, including linear kernels, the radial basis function (RBF) kernel, polynomial kernels, as well as orthogonal polynomial kernels (e.g., Chebyshev, Legendre, Laguerre, and Hermite polynomials) (Zhou et al., 2007). We utilize fuzzy SVMs coded in Matlab to predict sentiment for each product attribute in individual sentences. Specifically, a 10-fold cross-validation strategy is used to tune the parameters, in which a minimization procedure is used with the optimization tool in Matlab. In order to avoid local minima, 10 different initial values are randomly generated for the minimization process to obtain the possible global minima.

Table 5.2 Sentiment prediction results

Kernel function	Precision	Recall	F-score	#Support vector
Linear kernel	71.6%	71.7%	71.7%	1137
RBF kernel	73.6%	73.7%	73.7%	1134
Polynomial kernel ($d = 4$)	69.4%	69.4%	69.4%	513
4-item Chebyshev	73.0%	74.7%	73.8%	637
4-item Legendre	73.9%	75.4%	74.7%	648
4-item Laguerre	71.6%	75.8%	73.7%	664
4-item Hermite	75.1%	75.2%	**75.1%**	663
Average	72.6%	73.7%	73.2%	-

Table 5.3 Confusion matrix of best performance by 4-item Hermite SVM

		Predicted			
		Positive	Negative	Total	Recall
Actual	Positive	990	274	1264	78.3%
	Negative	356	908	1264	71.8%
	Total	1346	1182	2488	75.2%
	Precision	73.6%	76.8%	75.1%	F-score =75.1%

A 10-fold cross-validation method is also adopted for prediction results, in which the numbers of sentences in both the negative sentiment and the positive sentiment are made equal in order to obtain unbiased results. Precision, recall, and F-score (see Zhou et al., 2011b) are reported in Table 5.2, respectively. The bold numbers represent the best result. First, the average values of precision, recall, and F-scores are 72.6%, 73.7%, and 73.2%, respectively, indicating the proposed method gains a medium accuracy. Among all SVMs, the one with the 4-item Hermite kernel outperforms other SVMs in terms of F-score, which is the harmonic mean of precision and recall. Although the SVM model with the polynomial kernel function has the least number of support vectors, it has worst accuracy. Note that the SVM model with the RBF kernel function also achieves comparable results with those of orthogonal polynomial kernels. However, it nearly doubles the number of the support vectors, thus requiring more memory and computational resources. Furthermore, the confusion matrix for the best performance by the 4-item Hermite SVM is presented in Table 5.3. It shows that there are more false negatives (356) than false positives (274), which leads to a low recall for the negative sentiment, compared with other measures. This is probably caused by the features described in 4), i.e., V_{min}, A, D. By visual inspection, I find that V_{min} is larger than zero in many entries, which lead to predicting the negative sentiment as the positive sentiment by mistake.

5.6 Attribute Extraction and Refinement

5.6.1 Attribute Extraction by Association Rule Mining

According to Hu and Liu (2004b), I first generate product attributes using association rule mining. In this work, I define an item set as frequent if it appears in more than 1% (minimum support) of review data. The Apriori algorithm (Agrawal et al., 1993) works in two steps in association rule mining. In the first step, it finds all the frequent item sets from a set of transaction satisfy a user-specified minimum support. In the second step, it generates rules from the discovered frequent item sets. For this task, I only need the first step, i.e., finding frequent item sets, which are candidate product attributes. Nevertheless, the attributes generated are often redundant, regardless of the pruning methods applied. Furthermore, Carenini et al. (2005) point out that the attributes extracted are not arranged in a meaningful way or no structured organization is available that is beneficial to customer needs elicitation. Therefore, I propose a set of predefined attributes in a hierarchical model and a similarity measure is used to cluster lexically similar terms, such as photo and picture, to augment the user-defined attributes with sentiment and preference information. Note use case (see Figure 5.8) extraction and refinement are achieved similarly with association rule mining and similarity matching. Hence only attribute extraction and refinement are described in details.

5.6.2 Attribute Refinement by Similarity Matching

Five measures of semantic relatedness between two words, w_i and w_j in Python that can make use of WordNet are selected for similarity matching and they are briefly described as follows (Budanitsky and Hirst, 2001):

(1) Hirst-St-Onge similarity: It measures how similar two word senses are, based on the shortest path that connects the senses in the is-a (hypernym/hyponym) taxonomy in WordNet. It is calculated as $sim_{HS}(w_i, w_j) = C_{HS} - len(w_i, w_j) - k_{HS} \times d_{HS}$, where $len(w_i, w_j)$ is the path length between w_i and w_j, d_{HS} is the number of direction changes in the path, and C_{HS} and k_{HS} are constants.

(2) Leacock-Chodorow similarity: It measures how similar two word senses are, based on the shortest path that connects the senses and the maximum depth D_{LC} in the is-a taxonomy. It is calculated as $sim_{LC}(w_i, w_j) = -\log\left(\frac{len(w_i, w_j)}{2D_{LC}}\right)$.

(3) Resnik similarity: It measures how similar two word senses are, based on the information content of the least common subsumer (i.e., most specific ancestor node). It combines both ontology and corpus because the information content is dependent on the corpus used and the specifics of how the information content is created. It is computed as $sim_{Res}(w_i, w_j) = -\log p_s\,(lso(w_i, w_j))$, where $lso(w_i, w_j)$ is the least common subsumer between w_i and w_j, and $p_s(w)$ is the probability of encountering an instance of a synonym set w in some specific corpus. In this research, Brown Corpus (Malmkjær, 2002) is used, which contains over one million words from 15 text categories.

(4) Jiang-Conrath similarity: It makes use of information content based on the conditional probability of encountering an instance of a child of a synonym set, given an instance of a parent synonym set. The relationship is given by the equation: $sim_{JC}(w_i, w_j) = \frac{1}{2\log p_s\left(lso(w_i, w_j)\right) - \log p_s(w_i) - \log p_s(w_j)}$.

(5) Lin similarity: It makes use of the same elements in Jiang-Conrath similarity, but in a different form: $sim_L(w_i, w_j) = \frac{2\log p_s\left(lso(w_i,w_j)\right)}{\log p_s(w_i)+\log p_s(w_j)}$.

Taking two words, 'picture' and 'photo', as an example, I can calculate the similarity between them using the above-mentioned five measures, and the results are 0.333, 2.539, 5.562, 0.246, and 0.733, respectively. Since they are not in the same scale, each is normalized between 0 and 1 by dividing its maximum similarity scores.

For attributes with more than one word, I use $pa_i = [u_1, \dots, u_{m_1}]$ to denote a term attribute in the user predefined set, and $fa_j = [w_1, \dots, w_{m_2}]$ to denote a term attribute in frequent attributes mined through association rule mining. The similarity between them is calculated as the word average similarity (Carenini et al., 2005), i.e., $ave(pa_i, fa_j) = \frac{\sum_{i=1}^{m_1} max_j\left(ws(u_i,w_j)\right)+\sum_{j=1}^{m_2} max_i\left(ws(u_i,w_j)\right)}{2}$, where $max_j\left(ws(u_i, w_j)\right)$ is the maximum word similarity measure between u_i and the words in fa_j, and the five word similarity measures can be applied here.

First, in order to evaluate five similarity measures, a redundancy reduction measure is adopted as follows (Carenini et al., 2005):

$$RedundancyReduction = (N_{mfa} - N_{mpa})/N_{fa}, \tag{5.3}$$

where N_{mfa} is the number of matched attributes in the frequent attributes discovered by association rule mining, N_{fa} is the number of frequent attributes, and N_{mpa} is the number of matched attributes in the user predefined set. Note that the attributes that have exactly the same words and/or terms do not count in the calculation of redundancy reduction. Through association rule mining, I identify 116 product attributes and attribute levels for the Kindle Fire HD 7-inch model. Table 5.4 shows the results of redundancy reduction with different empirical thresholds. Here the threshold is defined as a function of the standard deviation of similarity scores. Note that this measure penalizes itself by increasing the value of N_{mpa} when a low threshold is used. Of all the five similarity measures, the similarity measures of Jiang-Conrath and Lin persistently perform better than other measures, and it seems that more redundancy reduction is obtained with higher thresholds (see Table 5.4).

Second, in order evaluate the classification accuracy, I first manually pick the typical product attributes from the extracted attributes to form the seed categories. Each category picks out three words or phrases to describe the product attributes and attribute levels that span the semantic space as large as possible based on our understanding. Then the five similarity measures with their respective optimal thresholds (i.e., corresponding to those in bold in Table 5.4) are used to classify product attributes based on the word's semantic similarity. The ground truth is built on manual sorting with three graduate students in engineering design with a strategy of majority voting. It generates 13 categories (see Figure 5.4) with one general category, i.e., Kindle Fire HD 7". Then the average prediction accuracy in terms of F-score (see Appendix A for its definition) is also shown in Table 5.4.

Table 5.4 Reduction of product attribute redundancy with different thresholds and classification accuracy

Similarity	Threshold					F-score
	0.2 std	0.4 std	0.6 std	0.8 std	1 std	
Hirst-St-Onge	0.218	0.245	0.303	**0.337**	0.342	68.5%
Leacock-Chodorow	0.226	0.270	0.292	0.321	**0.337**	65.4%
Resnik	0.224	0.279	0.352	0.359	**0.363**	71.2%
Jiang-Conrath	0.279	0.287	0.344	0.371	**0.372**	70.8%
Lin	0.273	0.315	0.362	0.372	**0.410**	**75.8%**

5.7 Latent Customer Needs Elicitation

5.7.1 Summarizing Customer Opinions on Product Attributes

Based on the sentiment analysis of individual product attributes, I am able to associate customer opinions with these product attributes in terms of their positive reviews versus negative reviews. First, based on the attribute extraction and refinement process, 13 major product attributes (see Figure 5.4) are identified, including screen, video, audio, reader, connectivity, online service, battery, dimension, customer service, price, storage, interface, and camera. This forms the attribute set, i.e., $A = \{a_k\}_{k=1}^{K}$, where $K = 13$. Furthermore, I also include the general review of the product itself as an attribute shown as KFHD 7" in Figure 5.4. The statistical summarization about the percentages of positive reviews and negative reviews of individual attributes is presented in Figure 5.4. First, more positive reviews are received than negative reviews as a general comment on the product itself from the attribute KFHD 7”. Second, screen, video, audio, reader, connectivity, dimensions, and prices have more positive opinions than negative opinions, showing that generally customers are satisfied with the current design of these attributes, especially audio and price. Third, customers complain disappointedly of product attributes, including camera and battery, whereas their opinions about storage, interface, online service, and customer service are mixed.

In order to have a clear understanding of customer preferences, I further identify attribute levels from attribute extraction and refinement, as shown in Figure 5.5. They form the set of attribute levels $A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$. For example, for the third attribute in Figure 5.5, $a_7 = storage$, and it has two levels, i.e., $a_{71}^* = 16\text{GB}$ and $a_{72}^* = 32\text{GB}$. Based on the association between the product attributes and their opinion polarity traced back from the sentences both reside in. I summarize their percentages of positive reviews vs. negative reviews within individual product attributes and their levels. Such information gives more insight into what customers like and dislike. For example, for storage, most customers complained of the 16GB version rather than the 32GB version, indicating that 16GB of storage capacity is not sufficient for most of the customers. For another example, despite the fact that dimension received a large portion of positive reviews (consistent with the reviews on weight), there are still about 1/3 of the customers not satisfied with the size. Furthermore, based on the comment information, I can rank them in terms of their frequency, as shown in Figure 5.6. By removing the general attribute, i.e., KFHD 7”, I recalculate their frequency and find that price, reader, screen, and video rank the top four. This shows the importance of these product attributes to a large extent (Rai, 2012).

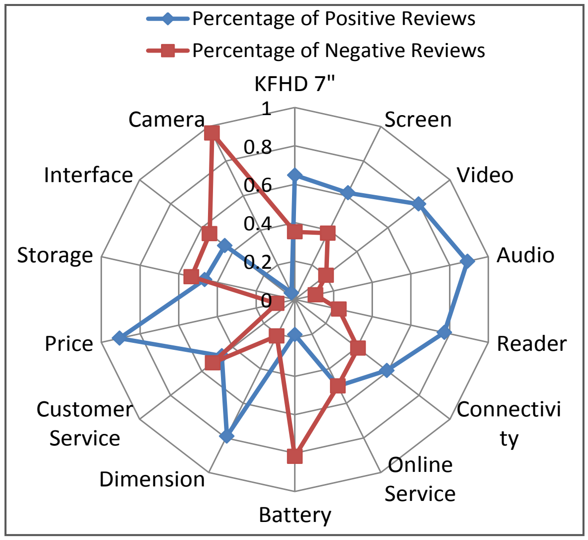


Figure 5.4 Customer opinions on individual product attributes

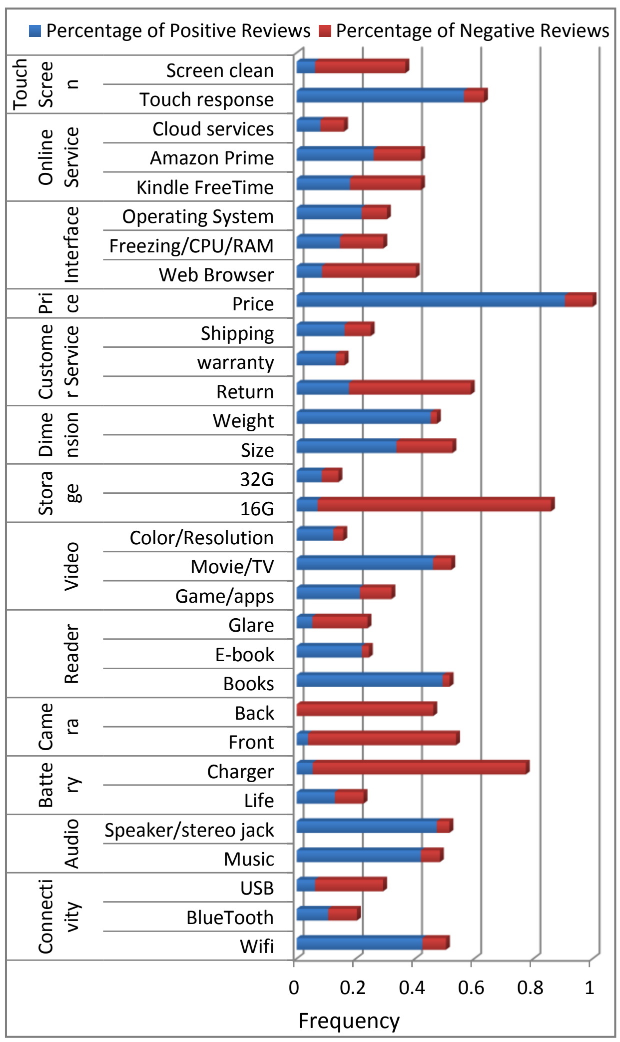


Figure 5.5 Customer opinions on attribute levels

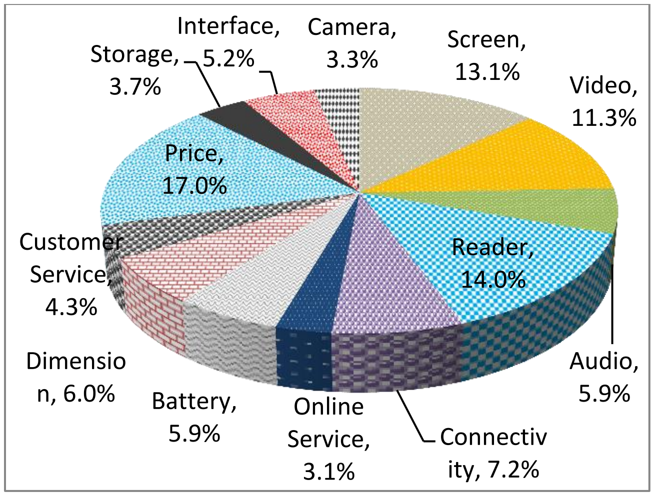


Figure 5.6 Frequency of product attributes in customer reviews

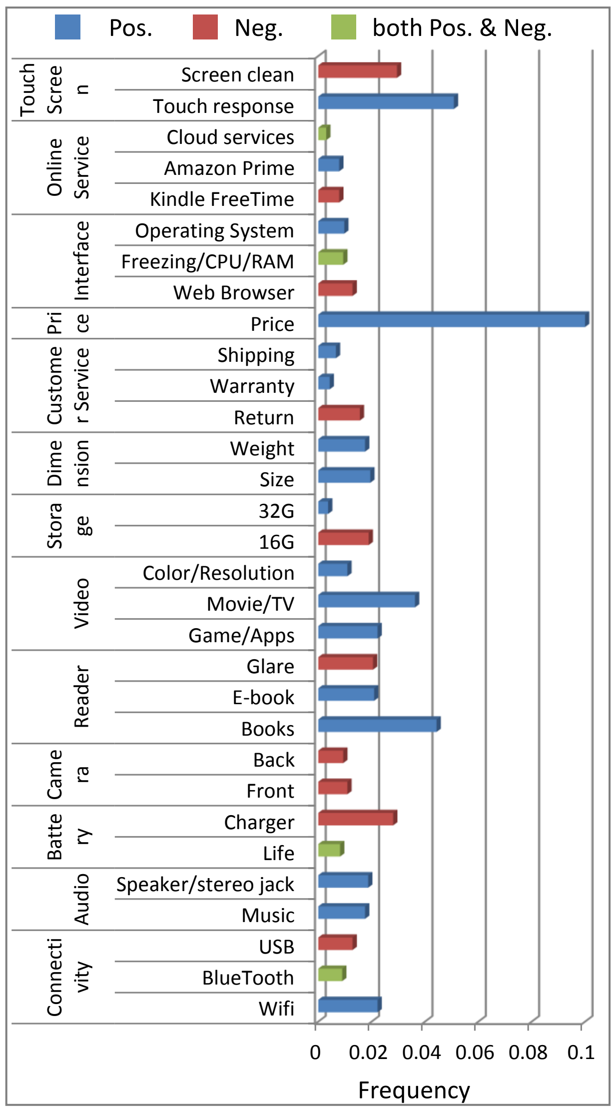


Figure 5.7 Frequency of attribute levels

5.7.2 Translating Customer Opinions into Customer Needs for Ordinary Cases

I divide the use cases into three interaction elements, including user types, interaction environments, and interaction contexts. Based on the association rule mining in sentiment analysis, I extract use cases as illustrated in Figure 5.8 based on the frequency commented on. Among them, I assume that typical adults, indoor with day light and seated form the ordinary interaction use case. Other cases with relatively low frequency are considered as extraordinary cases as shown in red rectangular in Figure 5.8. The figure demonstrates that extraordinary use cases consist of relatively rarer interaction elements compared with ordinary use cases. For example, kids/students are a combined set of children, nephews, niece, grandsons, granddaughters, sons, daughters, and kids mined from the reviews, and they account for 26.5% of all the user types. I translate customer opinions into needs for the ordinary use cases, based on which CBR is used to elicit latent customer needs for the extraordinary cases.

Table 5.5 Examples of how customer opinions are translated into customer needs for the ordinary use case

Sentiment	Customer review	Interpreted needs	Attributes
Positive	1) It’s so convenient that I can carry the kindle wherever I want in the house.	The tablet is portable. The tablet has good connectivity.	Dimension/ connectivity
	2) The kindle is easy to read and easy to use and see.	The user interface is easy to use and see.	User interface
	3) I purchased the Kindle for reading cheaper e-books. I love it.	The kindle support a variety of e-books with lower costs.	Reader
Negative	4) It’s fully charged and after I use it for a while it is still with a good amount of battery left. Yet when I put it in sleep mode, it’s dead in a few hours.	The battery life is long enough.	Battery
	5) The Kindle is kinda heavy when I hold it for a long time, reading, watching movies, or playing games.	The tablet has a support stand or is light enough to hold for reading, watching movies, or playing games for a long time.	Dimension
	6) I am disappointed that it has no back facing cam[era].	The tablet has a back facing camera.	Camera

First, sentiment analysis groups all the customer reviews about one product attribute or attribute level together in terms of positive and negative sentiments. This information facilitates transformation from customer opinions into customer needs. One advantage is that customer needs can be expressed as an attribute of the product, which ensures consistency and in turn facilitates subsequent translation into product specifications (Ulrich and Eppinger, 2003). As an example, I show some customer reviews, and how they are translated into customer needs in Table 5.5 for the ordinary case: typical adult, indoor with day light, and seated. From the positive customer reviews, the interpreted needs are expressed using positive phrasing with different levels of details that are as close to the raw data as possible. For example, the second review, the customer commented, “The kindle is easy to read and easy to use and see”. This directly is translated into “The user interface is easy to use and see”. For the negative customer reviews, often the opposite meaning will be the interpreted needs. As a rule of thumb, they are also expressed using positive phrasing with different levels of details. For example, the interpreted need for the last review, “I am disappointed that it has no back facing cam[era]”, is “The tablet has a back facing camera”.

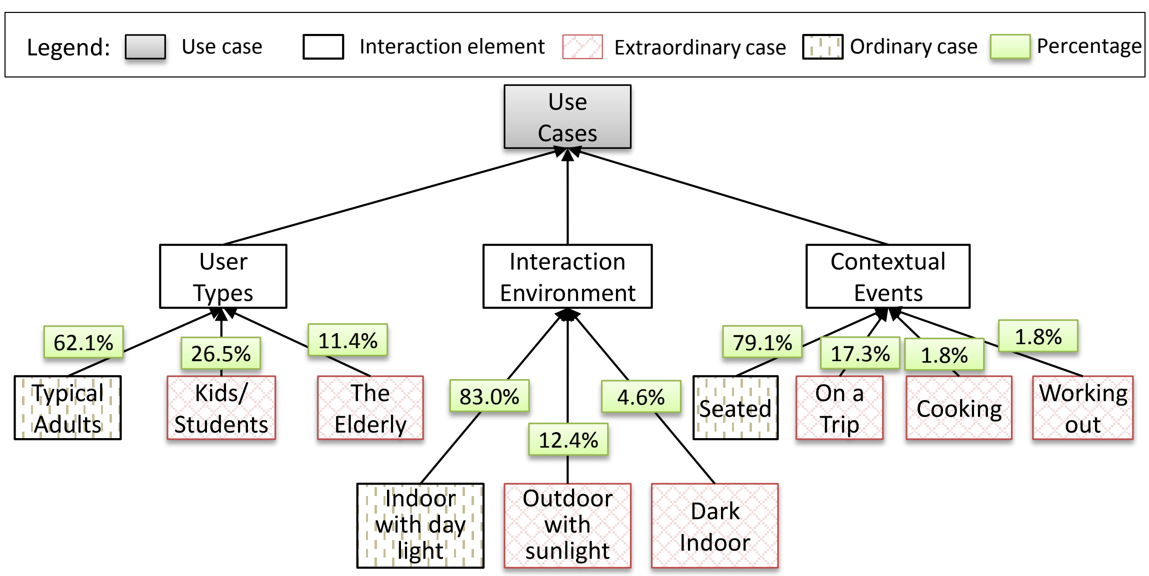


Figure 5.8 Extracted use cases from online user-generated product reviews

5.7.3 Eliciting Latent Customer Needs with Use Case Analogical Reasoning

We propose CBR for use case analogical reasoning. It employs a hybrid reasoning method by combing case-based and rule-based reasoning for case understanding to elicit latent customer needs (Zhou et al., 2011a). High-level customer needs are elicited first based on CBR and then a customized knowledge model compatible with rule-based reasoning is utilized to adapt an ordinary use case for an extraordinary use case. It has three major modules, including case database organization, case retrieval, and case adaptation.

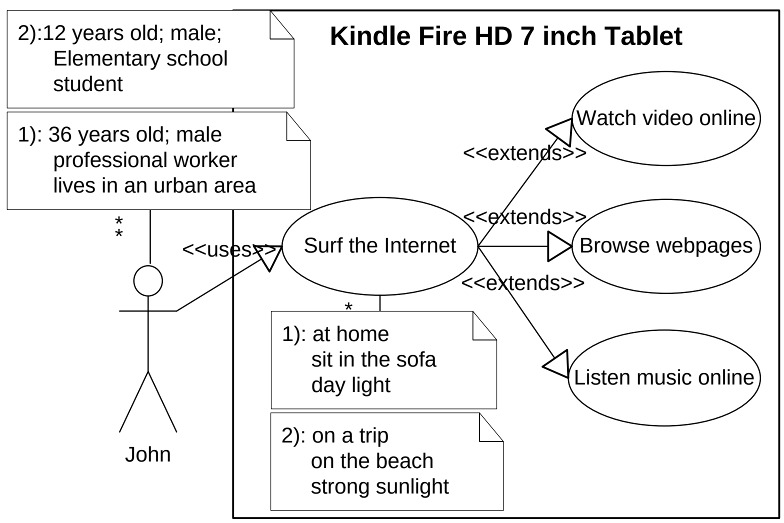


Figure 5.9 Case representation of 1) an ordinary case and 2) an extraordinary case

(1) Case database organization: The case database is denoted as $C^R = < C_1^R, \ldots, C_{M_R}^R, Ind, R_d >$, where M_R is the total number of refined cases for the time being, $Ind = < Activity, User, Ent, Evt >$ is a case index model, and R_d is a domain customized knowledge model in terms of rules for case adaptation. Cases are organized based on the case indexing model hierarchically. They are first grouped based on the use class $Activity$, followed by the user type $User$, the interaction environment Ent, and finally by the contextual event Evt. Thus, each case is represented in a top-down fashion. Use cases shown in Figure 5.8 can be constructed in terms of use case diagram using UML (unified modeling language), resulting in corresponding cases in the case database. As an example, I construct an ordinary use case and an extraordinary use case in Figure 5.9 as a form of case representation. In the extraordinary case, it describes that a boy

$(User)$ is surfing the Internet $(Activity)$ on a trip (Evt), on the beach with strong sunlight (Ent). Whenever such a case is identified, I need to retrieve the most similar case to reuse the customer needs and adapt them with R_d to elicit latent customer needs.

(2) Case retrieval: Case retrieval is the process of finding prior solved cases that are closest to the current case. Here solved cases mean that the customer needs are elicited for those cases. Given the extraordinary case C^e, the retrieval process proceeds with the following pseudo algorithm in Figure 5.10.

```
for i = 1: Activity  % Number of total activities in the case
  if Activity_i == Activity_e  % if the i-th case activity matches that of C^e
    for j = 1:User  % Number of total user types
      if User_j == User_e  % if the j-th user type matches that of C^e
        for k = 1: Ent  % Number of total interaction environment
          if Ent_k == Ent_e  % if the k-th interaction environment matches that of C^e
            for l = 1:Evt  % Number of total contextual event
              if Evt_l ==Evt_e  % if the l-th contextual event matches that of C^e
                Retrieve case C^ijkl
              else
                Calculate similarity between C^e & C^ijk*; Retrieve the one with maximum similarity
              end
            end
          else
            Calculate similarity between C^e & C^ij**; Retrieve the one with maximum similarity
          end
        end
      else
        Calculate similarity between C^e & C^i***; Retrieve the one with maximum similarity
      end
    end
  else
    NO similar case is retrieved; Manually elicit customer needs
  end
end
```

Figure 5.10 Case retrieval pseudo algorithm

It searches from the use class activity to check whether there is any case in the database that matches that of C^e; if so, it will go further to check whether the case user type matches that of C^e. It continues until it checks the interaction environment and the contextual event. If one case matches all the four variables, then the algorithm calculates the similarity between C^e and the retrieved case C^{ijkl}. If it only matches the first three variables, then the algorithm calculates the similarity between C^e and C^{ijk*}, where C^{ijk*} indicates a case with any contextual event. The same pattern applies to C^{ij**} and C^{i***}.

In such situations, I need to determine which one is the most similar case. Hence, a similarity measure is designed. We assume that the four variables (i.e., $Activity, User, Ent, Evt$) of two cases C_i and C_j are associated with m_c characteristics (e.g., age, gender for user type), i.e., $C_i = (c_1^i, c_2^i, \dots, c_{m_c}^i)$ and $C_j = (c_1^j, c_2^j, \dots, c_{m_c}^j)$. If the *i*-th characteristic is nominal, then $C(c_i^i, c_i^j) = 1$ when $c_i^i = c_i^j$, and 0 otherwise. If it assumes a real value or integer number, then $C(c_i^i, c_i^j) = |c_i^i - c_i^j| / (max_i - min_i)$, where max_i and min_i are the respective maximum and minimum values of the *i*-th characteristic. Then the similarity between C_i and C_j can be denoted as:

$$S_{ij}^C = \sum_{k=1}^{m_c} C(c_k^i, c_k^j) / m_c. \tag{5.4}$$

If no similar case is retrieved, I may have to manually elicit the customer needs for the given extraordinary case, and then save it in the case database for future use.

(3) Case adaptation: The domain customized knowledge model R_d adapts new cases to elicit latent customer needs. This is implemented by integrating the substitution and rule-based method into a soft reasoning mechanism (Pal and Shiu, 2004), including the following steps:

1) Substitution: It replaces invalid parts of the old use case with new content, according to key differences of a new case from the old one;

2) Rule-based adaptation: The system further refines the solution according to the customized knowledge model R_d ;

3) Evaluation: The design engineer performs evaluation and feedback for improvement;

4) Storage: If the adaptation is successful, the new use case, along with adaptation knowledge, is stored for future use. The customized knowledge model R_d is also updated if necessary.

As an example, I will use the extraordinary use case in Figure 5.9 as the new case, and the ordinary use case in Figure 5.9 as the case retrieved. The customer needs in Table 5.5 are translated into latent customer needs by case adaptation as shown in Table 5.6. First, the interpreted needs are usually assumed for the ordinary use case so that they are conditioned with square brackets (see the first column in Table 5.6) to facilitate case adaptation. Then the key differences between the ordinary use case and the extraordinary use case are highlighted. Therefore, the inferred latent customer needs are obtained by substitution in square brackets (see the second column of Table 5.6). The customized knowledge model R_d is simple and intuitive. Typical IF-THEN rules (see Appendix B for the whole list) include

1) IF the user type is kids/students, THEN the need is conditioned with parental-control or is children appropriate,

2) IF the interaction environment is outdoor with sunlight, THEN the need is conditioned with outdoor with sunlight, and

3) IF the contextual event is a trip, THEN the need is conditioned with trip characteristics (e.g., no WIFI connection, no power available).

According to these IF-THEN rules, the customized latent needs are shown in parentheses in the second column of Table 5.6. Finally, design engineers will examine the adaptation for refinement and improvement. If it is successful, the new case will be stored in the case database for future use. In such a way, the case database will grow progressively. Finally, the customer needs will be summarized with regard to each product attribute, which form the set that I am looking for $C_k^N = \left\{a_k, \{c_{kj}^n\}_{j=1}^{J_k}\right\}_{k=1}^{K}$.

Table 5.6 Example of latent needs elicitation with case adaptation

Interpreted needs	Latent needs
1) The tablet is portable [indoors for adults]. 2) The tablet has good connectivity [indoors].	1) The tablet is portable [outdoors for children]. 2) The tablet has good connectivity (on a trip) and [outdoors].
3) The user interface is easy to use and see [indoors for adults indoors].	3) The user interface is easy to use and see [outdoors for children].
4) The kindle support a variety of e-books with lower costs [indoors for adults].	4a) The kindle support a variety of e-books [text-books] (that are children-appropriate or with parental control) with lower costs. 4b) The kindle support reading [outdoors] (with sunlight).
5) The battery life is long enough.	5) The battery life is long enough (e.g., for long trips).
6) The tablet has a support stand or is light enough to hold [for adults] for reading, watching movies, or playing games for a long time.	6) The tablet has a support stand or is light enough to hold [for children] for reading, watching movies, or playing games (that are children-appropriate or with parental control) for a long time.
7) The tablet has a back facing camera→ The tablet takes picture easily [indoors].	7) The tablet takes picture easily [outdoors with sunlight].

5.8 Discussions

(1) Online user-generated product reviews: Evidence has shown that these reviews have become an important information venue for purchase decisions for customers and customer needs elicitation for designers. The proposed method in this chapter extracts reviews for one product from Amazon.com, which gives us opportunities to elicit customer needs for this product. Liu et al. (2013) point out that longer reviews receive a unanimously better evaluation than shorter reviews in terms of helpfulness for designers. The reason they found is that a long review covers customer preferences, mentions many different product attributes, and includes likes and dislikes of the product. Hence, these long reviews tend to be diagnostic in that they not only help designers in understanding and evaluating the quality and performance of products sold online, but also lead to a profound understanding of product use in different cases that often contribute to latent customer needs elicitation. However, shorter ones did not mention anything good or bad about product attributes or no information about the performance. From sentiment analysis point of view, those longer reviews with clear opinions are not only helpful for sentiment prediction, but also for customer preference elicitation.

In our study, about 38.7% of the reviews comment on the product itself without pointing out specific product attributes. Therefore, I exclude these reviews for customer needs elicitation though the preference information about the product shows the general popularity of the product (see Figure 5.4). I assume that the product reviews cover all of the possible product attributes. However, association rule mining only recovers those with frequent comments. Therefore, it is possible that product attributes with a very small number of comments are omitted in the study. However, on the one hand, these product attributes may be obsolete compared to popular ones (Tucker and Kim, 2011). On the other hand, they may be considered as must-be attributes in Kano model that are taken for granted when fulfilled, but result in dissatisfaction when not fulfilled (Yang, 2013). In this study, I mainly rank product attributes based on a simple frequency measure and their respective positive and negative opinions. Although simple, this measure is generally consistent with sophisticated importance measure, such as the review appearance rate measure and the local global normalization measure (Rai, 2012).

Therefore, it seems that many repetitive reviews of certain product attributes show their importance to a large extent.

(2) Sentiment analysis: First, the proposed method makes use of both affective lexicons and a supervised learning method, i.e., fuzzy SVMs, for sentiment analysis. The affective lexicons are constructed based on ANEW and WordNet. The seed list is rated with valence, arousal, and dominance, based on the three-dimensional model of emotions. It has been proved that this model is able to measure emotional reactions to stimuli in different contexts (Bradley and Lang, 1994). The seeds are then used to expand the affective lexicons using WordNet. Such a list of affective lexicons is domain independent and can be applied to predicting sentiments about reviews of different products.

Furthermore, in order to improve the prediction results, I propose fuzzy SVMs with different kernel functions. Using simple features extracted from affective lexicons, the model can achieve an F-score of 75.1%. As mentioned earlier, one limitation is that fuzzy SVMs are a supervised learning method, which needs a manual labeling process to create a ground truth for training and testing. In this research, I capitalize on the user-provided stars (i.e., rating) in reviews from Amazon.com. Those with 1 to 2 stars are considered as negative and those with 4 to 5 stars are considered as positive. However, this kind of rating information is not always available and reliable. Even a review with 1 to 2 stars, it is still likely that some product attributes are positively reviewed, and vice versa. A possible way to alleviate this problem is to manually label a relatively small number of reviews, and then a bootstrapping strategy can be used to alleviate the laborious labeling and training process for future work (Chen et al., 2012). Therefore, the ground truth is not optimal in this sense and it deteriorates the prediction performance to some degree.

Second, based on the previous research (Hu and Liu, 2004b), I extract product attributes based on association rule mining. However, due to the redundancy among the extracted attributes, a similarity matching method is proposed to refine the extracted product attributes. The shortcoming of this method is that attribute levels are still difficult to obtain or hierarchically organized with product attributes without human involvement. However, this is still an open question, and more research is still needed in the future.

(3) Latent customer needs elicitation: Latent customer needs can lead to major innovations. Unfortunately, many traditional elicitation techniques are unable to identify latent customer needs. As an alternative to traditional methods, I propose to make use of CBR for use case analogical reasoning that reuses customer needs from ordinary use cases and adapts them with a domain customized knowledge model for extraordinary use cases. Consistent with previous studies (Chen et al., 2013a; Hannukainen and Hölttä-Otto, 2006; Lin and Seepersad, 2007), this idea is inspired by transforming ordinary users into 'lead users' through transforming the ordinary use cases into infrequent, extraordinary use cases, including user types, interaction environments, and contextual events. The needs experienced in such use cases are thus considered as latent ones in this study.

However, unlike these previous studies, the proposed method is based on use case analogical reasoning without directly interviewing users. This is possible because the results from sentiment analysis provide us knowledge to build a case database for case reuse. Then, CBR retrieves the most similar case whenever an extraordinary use case is identified. The adaptation process is implemented by identifying reusable component in the source domain and analogical mapping with substitution in the target domain. Such a process is further enhanced with a domain knowledge model in terms of IF-THEN rule reasoning.

This process greatly improves the latent customer needs elicitation process and reduces designers' mental workload. From the designer's point of view, compared to traditional methods of data collection and customer needs elicitation, the proposed method can save a substantial amount of time and cost. Nevertheless, some precautions need to be taken. First, the reuse behavior may be difficult when there are only a small number of cases in the database. However, CBR is capable of learning new cases and can progressively increase its database. To finalize latent customer needs, effective evaluation strategies are needed and may be influenced directly by reuse behavior and prior experience of expert designers. However, for inexperienced designers, this system offers an opportunity for them to learn requirements elicitation and analysis.

5.9 Summary

This chapter addresses the challenges of latent customer needs elicitation through use case analogical reasoning from sentiment analysis of user-generated online product reviews. This perspective supplements the traditional methods of customer needs elicitation, especially those of eliciting latent ones. The proposed method of sentiment analysis combines a list of affective lexicons and a supervised learning method, i.e., fuzzy SVM. The affective lexicons are first initiated by the words in ANEW as seeds with quantitative, affective measures, and are then propagated with WordNet by identifying their synonyms and antonyms. These lexical features and two other syntactic features are used for sentiment prediction based on SVMs. Product attributes are also extracted and refined with association rule mining and similarity matching, respectively. Based on the results from sentiment analysis, I can summarize customer opinions on individual product attributes. At the same time, both ordinary and extraordinary use cases are extracted from association rule mining and are refined by similarity matching. Such extraordinary use cases greatly facilitate the process of latent customer needs elicitation using CBR by reusing and customizing ordinary use cases. Compared with traditional methods of customer needs elicitation, the proposed method can automate data collection and linguistic analysis, facilitate translation from customer likes and dislikes to explicit and latent customer needs, and help organize and prioritize customer needs.

The product attributes and attribute levels identified can be the input for customer preference modeling and quantification, and also become the choice set of viral product attributes. This is possible because among the customer needs elicited, the latent customer needs that delight customers unexpectedly are much more likely to be shared among peers in the online social network. Correspondingly, the design attributes and their attribute levels are more likely to be viral.

CHAPTER 6

PROSPECT THEORETIC MODELING OF CUSTOMER PREFERENCE INCORPORATING SUBJECTIVE EXPERIENCES FOR PRODUCT CHOICE DECISION MAKING

Product choices based on customer preferences involves complex decision making in the face of uncertainty. While affective elements are well-known to influence human decision making, prevailing computational models for analyzing and simulating human perception on customer preferences are mainly cognition-based models. In order to incorporate subjective experiences, including both affective and cognitive factors in the decision making process, a preference evaluation function based on cumulative prospect theory (CPT) is proposed for three different affective states and two different types of products (affect-rich vs. affect-poor). In order to tackle multiple parameters involved in the preference evaluation function, a hierarchical Bayesian model is proposed with a technique called Markov chain Monte Carlo (MCMC). It successfully estimates parameters that represent different cognitive tendencies and affective influences for customers both at an individual level and at a group level by generating posterior probability density functions of the parameters to incorporate inherent uncertainty. An experiment with four hypotheses is designed to test the proposed model. By using ANOVA (Analysis of variance), I find that 1) anxious participants tend to be more risk-averse than those in joy and excitement, 2) joyful and excited participants tend to be more risk-seeking than those in anxiety in preference-related choice decision making, 3) all participants tend to be averse to negative preferences, and 4) participants tend to value by feeling for affect-rich products and value by calculation for affect-poor products. Furthermore, five different types of models can predict choice decision making between product configurations with around 80% of accuracy. In summary, the results explain affective-cognitive decision making behavior in the complex domain of preference-based product choices and thus illustrate the potential and feasibility of the proposed method.

6.1 Subjective Experiences on Choice Decision Making

Human perception on preferences originates from the evolution of customers' affective states triggered by stimuli (events) along a chain of executing cognitive tasks involved in the human-product interaction process (Picard, 1997). Such experience-based preferences have two essential aspects: affect and cognition. Engineering design traditionally emphasizes products' functional requirements, yet with limited consideration of customers' affective and cognitive preferences (Falcioni, 2008). Human affect plays a significant and useful role in human decision making (Brown, 2008). It is, therefore, imperative for design research to bring in "the human dimension".

In this chapter, mental processes, such as sensation, memory, attention, perception, and problem solving, refer to aspects of cognition. They are described as cognitive or the cognitive system (i.e., the "analytic system") are thought to be necessary to perform a cognitive task, such as decision making (Wickens and Hollands, 1999). The analytic system makes use of conscious, deliberate cognitive processes with various algorithms and normative rules which produce logical behavior and maximize expected utility (Kahneman, 2003). Thus, the operations of the cognitive system are typically slower, more effortful, and more likely deliberately controlled, e.g., the process of solving a mathematical question.

On the other hand, affect is an encompassing term, including emotions, feelings, moods, and evaluations. An affective state is often a transient emotion, such as fear of a situation, which influences decision making (Simon, 1982). The affective system, also known as the "experiential system," employs past experiences, emotion-related associations and intuitions for decision making (Kahneman, 2003). The operations of the affective system are often fast, automatic, effortless, and associative. Customers tend to be more susceptible to affect-rich (i.e., hedonic) products than affect-poor (i.e., utilitarian) products. Affect-rich products are those that allow the consumer to feel pleasure, fun, and enjoyment from buying and using them, whereas affect-poor products are purchased for their practical and functional uses (Khan et al., 2004). For example, shoppers often experience impulsive buying for affect-rich products (e.g., a favorite music album) rather than for affect-poor products (e.g., a computer software CD). In addition, the impulsive buying process is considered extraordinary, exciting, and spontaneous (Rook, 1987). I argue that the affective system and the cognitive system need to work collaboratively in order to make the best decisions (Slovic et al., 2004).

While affective elements and subjective experience are well-known to influence human decision-making, prevailing computational models for analyzing and simulating human perception and evaluation on preferences are mainly cognition-based models (Ahn, 2010). Expected utility theory assumes that humans make decisions based on a deliberate cost-benefit analysis (Kahneman, 2000). Recent models based on behavioral decision theories focus on cognitive errors and heuristics in human judgments, but still ignore the role of emotion in human decision making (Brandstätter et al., 2006). Such a single cognitive perspective is not optimal for analyzing human decisions towards preference-based product choices, in which customers' affective states experienced at the time of decision making often influence their perceptions and choices (Ahn and Picard, 2005). The intimate coupling of affective and cognitive decisions has driven recent consensus on the integration of emotion and cognition (Scherer et al., 2001). Several computational mechanisms have emerged, which treat cognition as a necessary antecedent to emotion (Gratch and Marsella, 2004). However, the computational realizations of affect-cognition integration have largely been pragmatic, and the link between core cognitive functions and emotions has yet to be fully explored (Marinier Iii et al., 2009).

In this chapter, I 1) propose a preference-based product choice decision making model based on CPT that is more accurate than expected utility theory, 2) incorporate the influence of affect in the preference model as an extension to the original CPT theory, 3) estimate parameters involved in the model as a way to test the hypotheses and accommodate both individual heterogeneity and group homogeneity, and 4) aggregate individual utilities obtained from such preference models with copulas to capture interdependence among them.

6.2 Customer Preference Model Based on CPT

6.2.1 Model Architecture

Given a set of product attributes obtained using the method in Chapter 5, denoted as $A = \{a_k\}_{k=1}^{K}$, where K is the total number of product attributes. These product attributes embody the key characteristics of a product or service system. Each product attribute may assume a number of levels, either discrete or continuous, $A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$, where L_k is the total number of levels of a_k. Customer preference to a product or service is a holistic impression resulting from complex cognitive and affective interactions with the product configurations formed by different product attribute levels (Zhou and Jiao, 2013b), U in the customer's mental space. While u_{kl} indicates a quantitative measure of preference for a specific product attribute level, i.e., part-worth utility, U is the holistic measure of preference for the entire design, i.e., holistic product utility. With regard to various attribute levels, it is important that customers are able to make decisions based on their perceived preference. Therefore, the problem of customer preference-based product choice in the context of affective-cognitive decision making is formulated:

Given product attributes $A = \{a_k\}_{k=1}^{K}$ and attribute levels $A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$;

Find the optimal configuration of the product;

Maximize holistic perceived preference for a product configuration, i.e., $U = C(P, \mathbf{p})$;

Subject to

1) Affective influence on choice decision making,
2) Cognitive influence on choice decision making, and
3) Uncertainty involved in choice decision making,

where $C(P, \mathbf{p})$ is an aggregation function that computes the holistic preference for a product configuration P formed by specific product attribute levels with their choice probability vector $\mathbf{p}$.

In order to answer the problem formulated above, I propose a customer preference model based on CPT. The architecture is illustrated in Figure 6.1. The model utilizes experiment data to project the shape of the CPT preference function and the weighting function in order to deal with future decision making. The model comprises four consecutive phases, namely the perceptual phase, the affective-cognitive reasoning phase, the learning phase, and the evaluation phase. It assumes that the decision-making process is influenced by the customer's affective states, cognitive tendencies, and risk attitudes when the decision making is about to happen. In the perceptual phase, the user identifies the product types, and estimates the reference point corresponding to neutral preference, i.e., indifference. In the affective-cognitive reasoning phase, the CPT-based value function and the weighting function are formulated to evaluate preferences for diverse product attribute levels. In the learning phase, shape parameters involved in the preference model are estimated based on a hierarchical Bayesian model. In the evaluation phase, a CPT-based value function is used for preference evaluation. These four phases are described in details below.

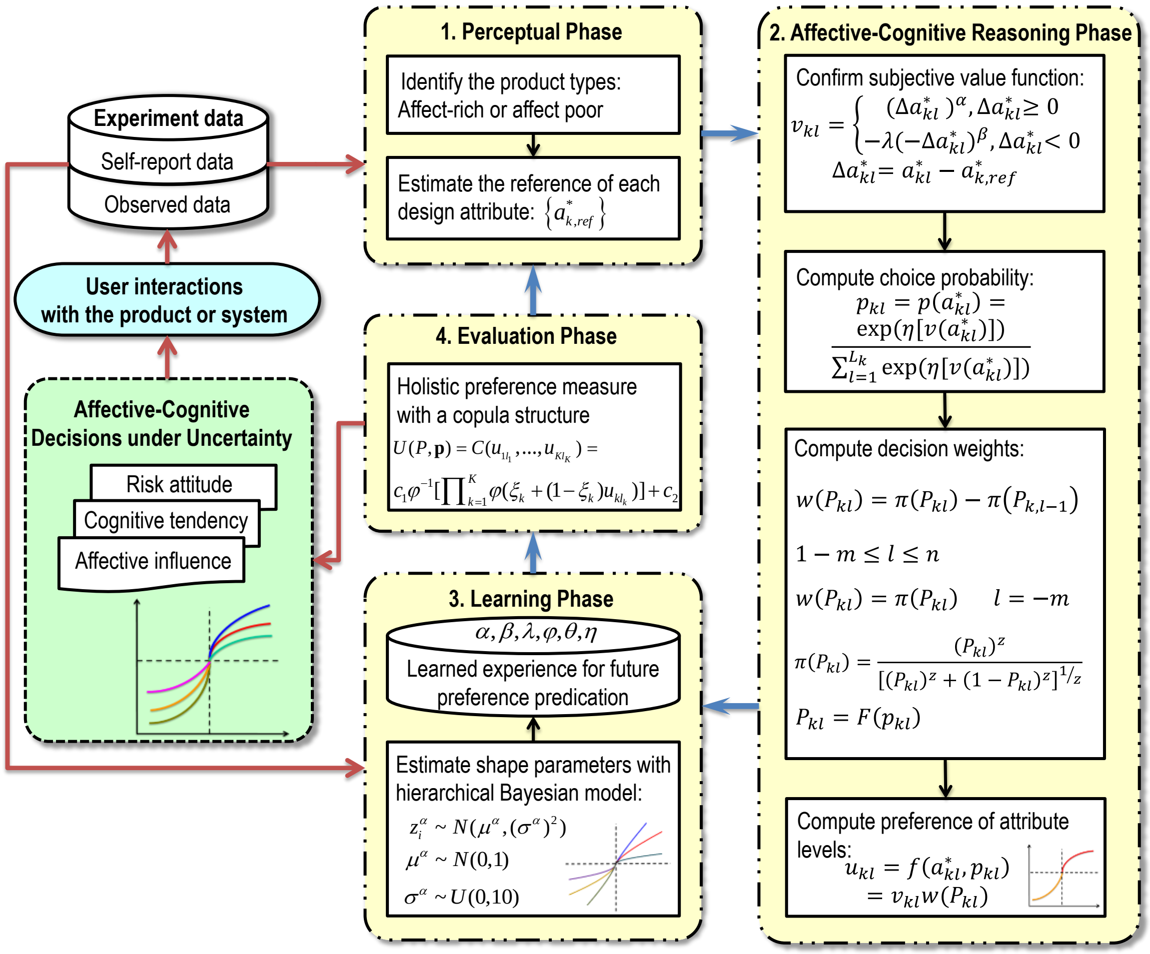


Figure 6.1 Preference model architecture based on cumulative prospect theory

6.2.2 Perceptual Phase

The customer's perceived preference of a product attribute level a^*_{kl} can be defined by a subjective value function, $v(a^*_{kl})$, weighted by his or her subjective probability. In the perceptual phase, the perceived preference of various options is identified relative to a certain product attribute level $a^*_{k,ref}$ that gives a neutral preference and acts as a reference point. Hassenzahl and Tracinsky (2004) point out that preference involves dynamic, context-dependent internal states of customers, including both instrumental and emotional aspects. It is hence likely that the reference point varies among different respondents. To hedge against this problem, I set up individual reference points for individual preference models to accommodate customer heterogeneity, and take a grand mean as the reference point for all the customers within one market segment to accommodate customer homogeneity.

6.2.3 Affective-Cognitive Reasoning Phase

(1) Subjective value function: CPT addresses important subjective influences on the preference-based product decision making process using a value function v as follows (Kahneman and Tversky, 1979):

$$v_{kl} = v(a_{kl}^*) = \begin{cases} (a_{kl}^* - a_{k,ref}^*)^{\alpha}, a_{kl}^* - a_{k,ref}^* \geq 0 \\ -\lambda(a_{k,ref}^* - a_{kl}^*)^{\beta}, a_{kl}^* - a_{k,ref}^* < 0' \end{cases} \tag{6.1}$$

where $a_{k,ref}^*$ is the reference point among attribute levels $\{a_{kl}^*\}_{L_k}$; The CPT subjective value function is defined with respect to a reference point and thus is reference dependent. In addition, α and β are parameters between 0 and 1, modulating the curvature of the subjective value function, which represents a decision maker's sensitivity to, risk attitude to, and affective influence on customer preferences. Four aspects below indicate how the value function accounts for the influence of affect and cognition on customer preferences (see Figure 6.2).

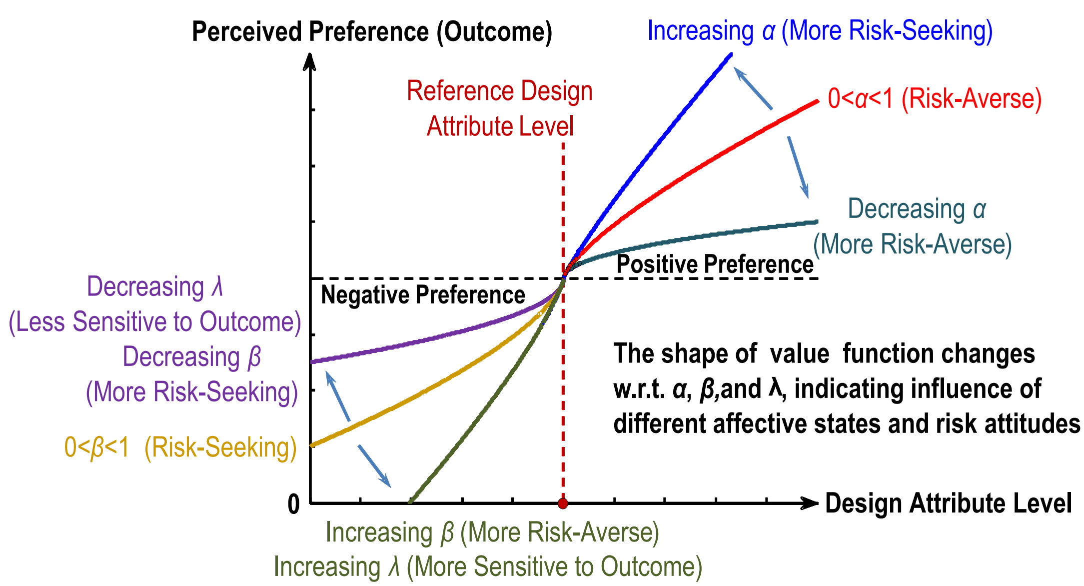


Figure 6.2 CPT-based preference value function

First, customers' *cognitive appraisal* of product attributes plays a significant role in assessing customers' perceived preferences. According to the appraisal theory, Ellsworth and Scherer (2003) state that customers evaluate product attributes (i.e., stimulus) in terms of perceived significance relative to needs and goals of the person concerned. Therefore, the more conducive the product is to achieving customers' goals and satisfying his/her needs, the more positive the perceived preference would be. When predicted emotions of future outcomes and/or genuine emotions elicited by the product attributes are very positive, impulsive buying may occur (Rook, 1987).

Second, the subjective value function has a *diminishing sensitivity* (i.e., $0 < \alpha, \beta < 1$), i.e., the change in preference decreases as the distance between the reference point and the target product attribute level increases (see Figure 6.2). Two psychological processes in constructing preferences are distinguished, i.e., valuation by feeling and valuation by calculation (Hsee and Rottenstreich, 2004). Valuation by feeling for affect-rich stimulus is scope-insensitive. For example, in Desvousges et al. (1993)'s study, the mean values of donation to save 2,000, 20,000, or 200,000 migrating birds from drowning in oil ponds are \$80, \$78, and \$88, respectively. This is explained by Kahneman et al. (1999) that Desvousges et al's questions evoke an affect-rich mental representation of an exhausted bird, its feathers soaked in black oil, unable to escape. The money the participants decide to donate is based on their affective reactions to this image. Valuation by calculation, however, is applicable to affect-poor stimuli and is more sensitive to scope, compared with affect-rich stimuli (Hsee and Rottenstreich, 2004).

Third, customers are *averse* to negative preferences. The inequality $\lambda > 1$ specifies the degree of aversion to negative preferences, meaning customers' perception on those attribute levels that are below the reference point, with larger values expressing more aversion and sensitivity to negative preferences (see Figure 6.2). The perceived preference of a product lies in a holistic impression. It is arguable that a negatively-perceived product attribute alone can jeopardize the positive preference towards the whole product, regardless of other attributes that are appealing to customers' perception (Zhou et al., 2010). This effect is more pronounced for affect-rich products than for affect-poor products. For example, Dhar and Wertenbroch (2000) report that participants are equally likely to choose a \$7 music CD (an affect-rich product) or a \$7 computer disk (an affect-poor product), but are five times as likely to surrender the computer disk if they are asked to give one up.

Fourth, on the one hand, studies on social-psychological and economic decisions have shown that customers with positive affective states tend to have an optimistic bias towards preferences, and therefore take greater risks (larger α) compared with neutral individuals (e.g., Ahn, 2010; Bracha and Brown, 2012). However, positive affective states promote risk-averse actions for negative preferences (larger β) in order to sustain their positive affective states (Isen, 2001). On the other hand, negative affective states influence customers' risk perception in different ways. For example, based on the appraisal-tendency framework (Lerner and Keltner, 2001), compared with neutral customers, anxiety and fear are related to situations of uncertainty and low control, which causes people to be risk-averse (smaller α and larger β), while anger is coupled with situations of certainty and high control, which triggers risk-seeking behavior (larger α and smaller β) (Gambetti and Giusberti, 2012; Lerner and Keltner, 2001).

(2) Choice probability: Original formulation of CPT is motivated for economic outcomes (gains or losses), and thus the choice behavior is crafted as a subjective probability by transforming the objective probability of an outcome using weighting functions (Tversky and Kahneman, 1992). It is true that different economic outcomes occur with varying probabilities. However, it is not the case for the product choice behavior, whereby product attribute levels are always available for customers to choose. Therefore, modeling of CPT choice probabilities should be consistent with the customer choice behavior.

Quantitative modeling to predict choice is an established area of research in marketing (Louviere et al., 2000) and product planning (Lewis et al., 2006). Using random utility discrete choice models, it is possible to predict customer preferences on different product attribute levels (Green and Krieger, 1985). The utility of a product attribute level a_{kl}^* to the customer is indicated by $v(a_{kl}^*)$. I can construct a closed form of choice probability adapted from the logit model (Train, 2003):

$$p_{kl} = p(a_{kl}^*) = \exp(\eta[v(a_{kl}^*)]) / \sum_{l=1}^{L_k} \exp(\eta[v(a_{kl}^*)]), \tag{6.2}$$

where $\eta > 0$ is a scaling parameter. As $\eta \to \infty$, the logit behaves like a deterministic model, while it becomes a uniform distribution as $\eta \to 0$.

(3) Weighting function: Using cumulative probabilities can lead to evaluations that choose first-order stochastically dominated choices rather than the dominating one (Park et al., 2013). I thus propose to formulate a cumulative prospect evaluation. An attribute assuming multiple levels, $A_k^* = \{a_{kl}^*\}_{L_k}, 1 \leq l \leq L_k$, can be transformed to $m + n + 1$ possible outcomes as perceived by a decision maker. Arrange these outcomes in an ascending order, i.e., $v_{-km} < \cdots v_{k0} < \cdots v_{kn}$, along with the respective choice probabilities, $p_{-km}, \ldots, p_{k0}, \ldots, p_{kn}$ or $\mathbf{p} = \{p_{kx}\}$. Note that v_{k0} corresponds to indifference at the reference level; those smaller than v_{k0} are related to the negative preferences of attribute levels; and those larger than v_{k0} are attributed to the positive preferences of attribute levels. The decision maker evaluates each attribute level with the associated choice probability, and thus the perceived preference for a_{kl}^* after probability distortion can be defined as:

$$u_{kl} = f(a_{kl}^*, p_{kl}) = v_{kl} w(P_{kl}) , \tag{6.3}$$

where $w(\cdot)$ is a decision weight, determined as the first-order difference of a probability weighting function applied to the cumulative probability $P_{kl} = F(p_{kl})$, that is, $w(P_{kl}) = \pi(P_{kl}) - \pi\left(P_{k,l-1}\right)$, if $1 - m \leq l \leq n$; and $w(P_{kl}) = \pi(P_{kl})$, if $l = -m$. I will adopt a particular probability-weighting function:

$$\pi(P_{kl}) = \frac{(P_{kl})^z}{[(P_{kl})^z + (1 - P_{kl})^z]^{1/z}}, \tag{6.4}$$

where $0 \leq z \leq 1$ specifies the curvature of the weighting function, such that $z = \delta$ stands for positive preferences (i.e., $w = w^+$) and $z = \theta$ suggests negative preferences (i.e., $w = w^-$). Decreasing the value of z makes the function become more curved. This function shows that customers tend to overweigh low probabilities with extreme preference outcomes and underestimate moderate and high probabilities (see Figure 6.3). One good example is that customers often overweigh the value of first-class cabins with a lower choice probability, but underestimate the value of economy cabins with a higher choice probability. This effect is more evident for affect-rich products than for affect-poor products. For example, Rottenstreich and Hsee (2001) show that participants are willing to pay $20 and $5 for a 1% chance to win a $500 gift certificate for a vacation to Europe (affect-rich) and school tuition (affect-poor), respectively. However, if the chance is 99%, participants are willing to pay \$450 and \$478 for the vacation and tuition, respectively.

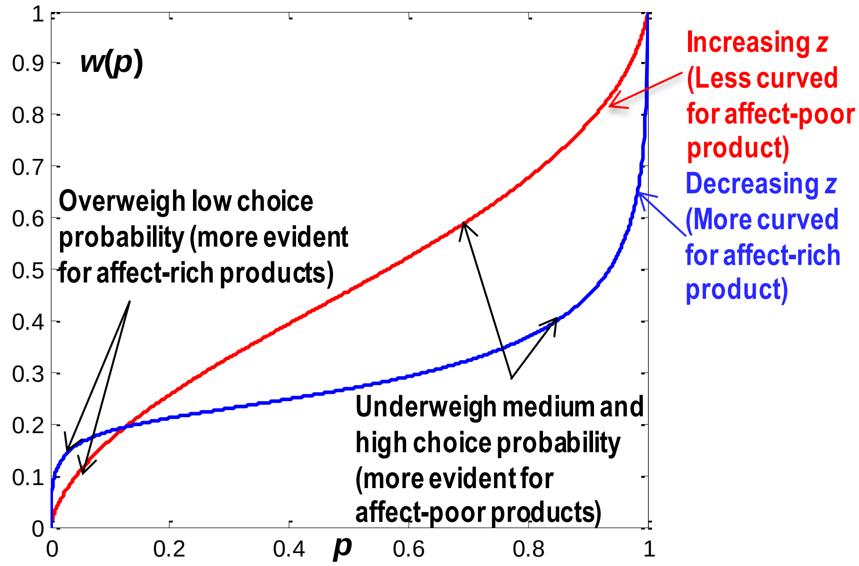


Figure 6.3 CPT-based weighting function for preference modeling

6.2.4 Learning Phase

In the learning phase, the parameters involved in the preference evaluation function will be estimated from the experimental data. Among many others, a hierarchical Bayesian method is utilized as elaborated below:

(1) Hierarchical representation of parameters: The hierarchical Bayesian representation of parameters involved in the CPT-based preference model is shown in Figure 6.4. $d_{ji} = \{1, 0\}$ indicates j-th participant makes the i-th choice about two alternative product configurations. If it is 1, the participant chooses configuration A and 0 configuration B, where $1 \leq i \leq N_d, 1 \leq j \leq M_p$, N_d is the total number of the decisions made for each participant, and M_p is the total number of participants. Whether the participant chooses configuration A or configuration B depends on the perceived preference calculated by Eq. (6.6). According to Nilsson et al. (2011), in order to account for inconsistencies for j-th participant involved in the i-th decision making (for $1 \leq i \leq N_d$), a logistic choice rule is introduced:

$$q_{ji}(A\ over\ B) = 1/(1 + \exp(\eta_j(C(A) - C(B)))), \tag{6.5}$$

where C is a copula function to calculate the holistic perceived preference to a product configuration (see Eq. (6.6)) and $\eta_j > 0$ is the sensitivity parameter. As $\eta_j \to \infty$, the logit behaves like a deterministic model to choose whichever with larger perceived preferences, and it becomes a uniform distribution as $\eta_j \to 0$. The data d_{ji} follow the Bernoulli distribution with parameter q_{ji}, i.e., $d_{ji} \sim Bernoulli(q_{ji})$.

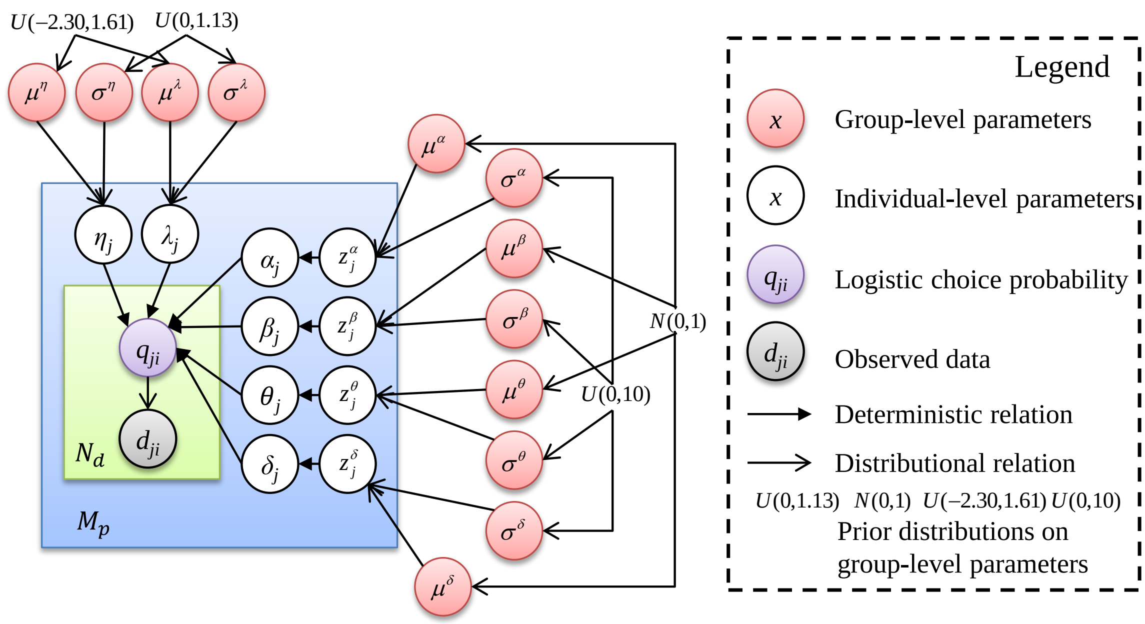


Figure 6.4 Hierarchical Bayesian parameter representation.

Note d_{ji} is the only observed decision making data, q_{ji} is the logistic choice probability to account for choice inconsistencies, and x indicates any parameters. Those in white circles are individual-level parameters, which are governed by group-level parameters in red circles, which are further governed by known prior distributions

For the j-th participant, it can be seen that q_{ji} is a function of parameters including $\eta_j, \alpha_j, \beta_j, \lambda_j, \delta_j$, and θ_j, which are shown in Figure 6.4 with arrow ' $\leftarrow$'. Of all the parameters, $\alpha_j, \beta_j, \delta_j$, and θ_j are between 0 and 1. Since these parameters will have the same treatment in the model, only the description of α_j is detailed here. According to Rouder and Lu (2005), a probit transform model is used as follows: Let Φ denote the standard normal cumulative distribution function, and I assume $\alpha_j = \Phi(z_j^\alpha)$, where $\alpha_j \in [0,1]$ and $z_j^\alpha \in R$. Following the probit transform model, I can have $z_j^\alpha = \Phi^{-1}(\alpha_j)$. Meanwhile, the probitized individual parameter is assumed to follow an independent normal distribution at the group level, i.e., $z_j^\alpha \sim N(\mu^\alpha, (\sigma^\alpha)^2)$ (denoted as '⟵'), where μ^α and σ^α are group-level parameters governing the distribution of individual-level parameters (i.e., z_j^α or α_j). Based on the three-sigma rule, z_j^α centers on zeros, but ranges mainly in $(\mu^\alpha - 3\sigma^\alpha, \mu^\alpha + 3\sigma^\alpha)$ with a probability of 99.73%. Thus, in order to effectively constrain the values of α_j between 0 and 1, the group-level parameters also need known prior distributions on them, showing the prior knowledge about these parameters. Based on the previous research (Gelman and Hill, 2007), for the mean, it follows a standard normal distribution, namely, $\mu^\alpha \sim N(0,1)$. For the standard deviation, it follows an uninformative uniform prior: $\sigma^\alpha \sim U(0,10)$. Using this manipulation, I am able to not only estimate individual-level parameters (e.g., α_j) to account for individual differences, but also group-level parameters (e.g., μ^α and σ^α) to explain customer homogeneity.

Parameters λ_j and η_j take positive values. According to Nilsson et al. (2011), their priors can have lognormal distributions, i.e., $\lambda_j \sim LN(\mu^\lambda, (\sigma^\lambda)^2)$ (η_j has the same treatment of λ_j) and the mean lies in an interval between 0.1 and 5, i.e., -2.30 and 1.61 on the natural log scale. Therefore, the mean at the group level follows $\mu^\lambda \sim U(-2.30,1.61)$, if an uninformative uniform prior distribution is assumed for the lognormal mean, the standard deviation is 1.13 (i.e., $\sqrt{1/12(2.30+1.61)^2} = 1.13$). Hence, it is reasonable that the standard deviation at the group level follows the uniform distribution: $\sigma^\lambda \sim U(0,1.13)$. Thus, Figure 6.4 is summarized as follows:

1) $d_{ji} \sim Bernoulli(q_{ji})$, where d_{ji} is the decision making data and is either 0 or 1;
2) $q_{ji}(A\ over\ B) = 1/(1+\exp(\eta_j(C(A) - C(B))))$, where C (Eq. (6.6)) is a copula function of to calculate the holistic utility of the product and is a function of parameters $\eta_j, \alpha_j, \beta_j, \lambda_j, \delta_j,$ and θ_j;
3) $\alpha_j, \beta_j, \lambda_j, \delta_j, \eta_j,$ and θ_j are individual-level parameters, among them $\alpha_j, \beta_j, \delta_j,$ and θ_j are between 0 and 1, and have the same treatment. Take α_j as an example, $\alpha_j = \Phi(z_j^\alpha)$, or $z_j^\alpha = \Phi^{-1}(\alpha_j)$. Thus, using the probit transform, z_j^α takes value from $(-\infty, \infty)$, and further $z_j^\alpha \sim N(\mu^\alpha, (\sigma^\alpha)^2)$; η_j and λ_j take positive values and have the same treatment. Take λ_j as an example, $\lambda_j \sim LN(\mu^\lambda, (\sigma^\lambda)^2)$,

and further $\mu^{\lambda} \sim U(-2.30,1.61)$ and $\sigma^{\lambda} \sim U(0,1.13)$;

4) $\mu^{\alpha}, \sigma^{\alpha}, \mu^{\beta}, \sigma^{\beta}, \mu^{\theta}, \sigma^{\theta}, \mu^{\delta}, \sigma^{\delta}, \mu^{\eta}, \sigma^{\eta}, \mu^{\lambda}, \sigma^{\lambda}$ are group level parameters, which govern individual parameter distributions. They also have prior distributions that govern them, i.e., $\mu^{\alpha}, \mu^{\beta}, \mu^{\theta}, \mu^{\delta} \sim N(0,1)$, $\sigma^{\alpha}, \sigma^{\beta}, \sigma^{\theta}, \sigma^{\delta} \sim U(0,10)$, $\mu^{\eta}, \mu^{\lambda} \sim U\ (-2.30, 1.61)$, and $\sigma^{\eta}, \sigma^{\lambda} \sim U(0,1.13)$.

(2) MCMC for parameter estimation: As mentioned previously, hierarchical Bayesian models construct the *posterior* distribution to estimate the parameters, i.e., $P(\alpha_j|d_j) \propto P(\alpha_j)P(d_j|\alpha_j)$. However, it is often the case that the posterior distribution is high-dimensional, complex, and unavailable in the closed form and, therefore, the analytical calculations involved are too difficult to perform (Lunn et al., 2000). The emergence of MCMC (Neal, 1993) has eliminated this analytic bottleneck. The idea behind MCMC is that it generates samples by constructing an ergodic Markov chain (i.e., irreducible and aperiodic, where irreducible means that every state is eventually reachable from any start, and aperiodic means that the chain does not get caught in cycles), which converges after a certain number of steps to the desired posterior probability distribution. These samples can then be used to approximate multidimensional integrals. Particularly, Gibbs sampling is used as a special case of the Metropolis-Hastings algorithm (Cowles and Carlin, 1996). The gist of Gibbs sampling is that given a multivariate distribution and some initial value for each parameter, it samples each parameter from the distribution of that parameter conditioned on the remaining parameters, making use of the most recent values and updating the parameter with its new value as soon as it has been sampled. This procedure is conducted recursively from the posterior conditional distribution until it converges.

I define individual parameters as a 6-tuple $\kappa_j = [\alpha_j, \beta_j, \delta_j, \theta_j, \eta_j, \lambda_j]$, and $\kappa = [\kappa_1, \dots, \kappa_M]$ as a $6 \cdot M$ tuple for M participants. Define $\mathrm{K} = [\mu^{\alpha}, \sigma^{\alpha}, \mu^{\beta}, \sigma^{\beta}, \mu^{\theta}, \sigma^{\theta}, \mu^{\delta}, \sigma^{\delta}, \mu^{\eta}, \sigma^{\eta}, \mu^{\lambda}, \sigma^{\lambda}]$, containing 12 group-level parameters. Assuming the parameters to estimate are denoted as $\Theta = [\vartheta_1, \dots, \vartheta_n] = [\kappa, \mathrm{K}]$, the Metropolis-Hasting sampling algorithm is summarized as follows (Cowles and Carlin, 1996):

1) Initialize $\vartheta^1 = [\vartheta_1^1, \dots, \vartheta_n^1]$;

2) for $i = 1, \dots, S_t$, where S_t is the total number of samples

for $j = 1, \dots, n_p$, where n_p is the number of parameters

draw a sample $u \sim U(0,1)$;

draw a candidate sample along the *j*-th direction of the proposal distribution $\vartheta^c = q(\vartheta^c|\vartheta_{-j}^{i+1}, \vartheta_j^i)$ without changing other coordinate values, where ϑ_{-j}^{i+1} denotes that all coordinate values of (*i*+1)-th sample are given except *j*-th coordinate value;

if $u < A(\vartheta^c, \vartheta_j^i) = \min(1, \frac{p(\vartheta^c|\vartheta_{-j}^{i+1})q(\vartheta_j^i|\vartheta^c, \vartheta_{-j}^{i+1})}{p(\vartheta_j^i|\vartheta_{-j}^{i+1})q(\vartheta^c|\vartheta_j^i, \vartheta_{-j}^{i+1})})$, then $\vartheta_j^{i+1} = \vartheta^c$;

else $\vartheta_j^{i+1} = \vartheta_j^i$.

The proposal distribution is often chosen as symmetric, such as $q(\vartheta^c|\vartheta_{-j}^{i+1}, \vartheta_j^i) = N(\vartheta_j^i, (\sigma_j^i)^2)$ so that $q(\vartheta_j^i|\vartheta^c, \vartheta_{-j}^{i+1})$ and $q(\vartheta^c|\vartheta_j^i, \vartheta_{-j}^{i+1})$ can cancel out. The proposal distribution for Gibbs sampling (Geman and Geman, 1984) is specially chosen as $q(\vartheta^c|\vartheta_{-j}^{i+1}, \vartheta_j^i) = p(\vartheta^c|\vartheta_{-j}^{i+1})$ so that the acceptance rate is always 1, which leads to fast convergence.

It can be seen that the most recent draws are actually dependent on the previous ones. If the dependence is not strong, then convergence can happen quickly. There are different ways to assess convergence, such as diagnostic tools (Cowles and Carlin, 1996). One informal, graphical method of assessing the degree of dependence is to plot the autocorrelation functions of the chains (Rouder and Lu, 2005), that is, the correlation between values of the process of the Markov chain of different times as a function of the time lag. One way to decrease autocorrelation is to thin the sample, using only every *k*-th sample. Furthermore, I run two Markov chains to ensure convergence and comparison. For a more detailed description, please refer to (Zhou and Jiao, 2013a).

6.2.5 Evaluation Phase

In order to aggregate individual preference evaluation functions (i.e., part-worth utility functions), considering the interdependence between individual product attributes, I propose to apply nested Archimedean copulas (A popular kind of copula that allows dependence modeling in arbitrarily high dimensions with only one parameter, governing the strength of dependence (see Zhou and Jiao, 2013b for details)). A brief description is given here. The holistic perceived preference of a design configuration of K product attributes $A = \{a_k\}_{k=1}^{K}$ with its l_k-th level is aggregated by an Archimedean copula, i.e.,

$$U = C\left(u_{1l_1}, \ldots, u_{KL_K}\right) = c_1 \varphi^{-1}\left[\prod_{k=1}^{K} \varphi\left(\xi_k + (1-\xi_k) u_{kl_k}\right)\right] + c_2, \tag{6.6}$$

where u_{kl_k} is between 0 and 1 by normalizing u_{kl} in Eq. (6.3), indicating preference of individual-attribute value, i.e., part-worth utilities, and U is the aggregated multi-attribute value, i.e., holistic product utility; $c_1 = 1/\left(1 - \varphi^{-1}\left(\prod_{k=1}^{K} \varphi(\xi_k)\right)\right)$, $c_2 = 1 - c_1$, ξ_k is a constant, $0 \leq \xi_k < 1$, and the generating function φ is (a) continuous on the domain $[0,1]$, (b) strictly increasing on the domain $[0,1]$, and (c) $\varphi(0) = 0$ and $\varphi(1) = 1$. It takes the form $\varphi(\xi_k) = \frac{1-\exp(\varsigma \xi_k)}{1-\exp(-\varsigma)}$, where $\varsigma \in R \backslash \{0\}$.

The perceived preference based on CPT involves various mental processes that are thought to drive decision making to optimize product choices (Nilsson et al., 2011). According to Eq. (6.6), under CPT, a preference prospect $U(P_i, \mathbf{p}_i)$, is preferred to another prospect, $U(P_j, \mathbf{p}_j)$, for a specific customer if and only if $U(P_i, \mathbf{p}_i) \succ U(P_j, \mathbf{p}_j)$, and is indifferent to another if $U(P_i, \mathbf{p}_i) \sim U(P_j, \mathbf{p}_j)$. Based on the descriptions above, preferences are determined jointly by a subjective value function that evaluates individual preference of specific product attribute level with regard to a reference point, and by the decision weights that capture an individual's distortion of choice probability. Furthermore, the shaping parameters embedded in the subjective value function capture customers' psychological processes, including risk attitudes, cognitive tendencies, and affective influences. Finally, a holistic measure of multi-attribute preference is obtained by aggregating individual-attribute preferences. Under the circumstances of product design for multiple product attributes $A = \{a_k\}_{k=1}^{K}$ and each with several levels $A_k^* = \{a_{kl}^*\}_{l=1}^{L_k}$, the cumulative prospect theoretic model can be used to evaluate customers' holistic preferences of alternative design configurations in the design space.

6.3 Empirical Study for Model Parameter Estimation and Validation

6.3.1 Background

The experiment focuses on an aircraft cabin interior design. It aims to create positive preferences and experience in the aircraft cabin. In order to disguise proprietary information, the aircraft cabin interior design is simplified. The product attributes and their attribute levels are given in Table 6.1. The aircraft cabin environment is different from others in many aspects. For example, passengers are potentially exposed to air contaminants, such as ozone, carbon monoxide, various organic chemicals, and biological agents. Air crew members operate in an environment that exposes them and passengers to hazardous vibration and noise levels, which influence their coordination and vision. The vibration during flight cannot be eliminated, but its effects can be minimized by practicing good prevention techniques. A sculpted ceiling with gentle curves makes an aircraft cabin feel more spacious and safer than hard edges. For example, it can be designed to hit the curved surfaces of the cabin in a way so that it makes the ambient light feel softer and the cabin feel much bigger. For the interior color of the cabin, the right color can create the desired mood and atmosphere. Blue can send a message of clean or fresh fragrance, while pink may imply something sweet. Colors also can influence a person's perception of humidity and temperature. For example, orange may make a person feel warmer, blue/green may make them cooler, and green (moist) and orange (dry) can signify extremes of humidity. Different combinations of product attributes are able to elicit different levels of customer preferences. Therefore, in order to provide proper product and service offerings in a profit-maximizing way, airliners need to understand cabin configurations from the perspectives of the customers and manage airline resources to deliver pleasant cabin customer experience while customers need to calculate the cost-benefit tradeoff for superior cabin experience. The key is to develop a decision making model that can evaluate customer preferences with regard to different cabin configurations.

6.3.2 Cabin Configurations

Given all the product attributes and their levels in Table 6.1, a total number of $5 \times 2 \times 3^8$ = 65610 combinations can be constructed. To overcome such an explosion of configurations, orthogonal product configurations are used. Twenty-seven orthogonal product configurations are generated based on the principle of design of optimal experiment (Nair et al., 1995) with SPSS 15.0 in Table 6.2, in which columns 2-11 indicate the specification of the product configurations.

Table 6.1 Product attributes and levels of aircraft cabin interior design

Product Attribute		Attribute Levels	
a_i	Description	a_{ik}^*	Description
a_1	Interior Color (right color will create the desired mood, atmosphere, and humidity)	a_{11}^*	Blue (cool, fresh fragrance)
		a_{12}^*	Green (moist, peace)
		a_{13}^*	Orange (dry, warm)
		a_{14}^*	Pink (sweet)
		a_{15}^*	White
a_2	Legroom	a_{21}^*	Restricted: seat pitch < 25”, width <17”
		a_{22}^*	Adequate: seat pitch 25”-30”, width 17” -20”
		a_{23}^*	Spacious: seat pitch > 30”, width > 20”
a_3	Noise: engine vibration-related noise	a_{31}^*	Low: below 84dB
		a_{32}^*	Medium: 84dB-89.9dB
		a_{33}^*	High: 90dB-95.9dB or above
a_4	Interior Light	a_{41}^*	Not adjustable
		a_{42}^*	Adjustable: basic brightness
		a_{43}^*	Premium adjustable: subtly change tone according to the time of day
a_5	In-flight Entertainment (IFE)	a_{51}^*	No in-flight entertainment
		a_{52}^*	Limited entertainment
		a_{53}^*	Rich entertainment
a_6	Interior Patterns	a_{61}^*	A sculpted ceiling with gentle curves
		a_{62}^*	Hard lines or flat surfaces
a_7	Cabin Air Pressure	a_{71}^*	Normal (close to the ground)
		a_{72}^*	Relatively low (altitude equivalent of 6,000 feet)
		a_{73}^*	Low (altitude equivalent of 8,000 feet)
a_8	Relative Humidity	a_{81}^*	Relatively low (20%-30%)
		a_{82}^*	Low (10%-20%)
		a_{83}^*	Extremely low (<10%)
a_9	Vibration	a_{91}^*	Strong: over 0.05m/s
		a_{92}^*	Medium: between 0.02-0.05m/s
		a_{93}^*	Weak: below 0.02m/s
a_{10}	Contaminant (Ozone, carbon monoxide, organic chemicals, and biological agents)	$a_{10,1}^*$	High: sterilize and clean the plane once a week
		$a_{10,2}^*$	Medium: sterilize and clean the plane twice a week
		$a_{10,3}^*$	Low: sterilize and clean the plane daily

Table 6.2 Aircraft cabin interior design configurations for evaluations

#Configuration	Interior Color	Legroom	Noise	Interior Light	IFE	Interior Pattern	Cabin air pressure	Relative humidity	Vibration	Contaminant
1	green	restricted	medium	basic adjustable	no	sculpted ceiling with gentle curves	low	20%-30%	weak	low
2	pink	restricted	high	not adjustable	limited	sculpted ceiling with gentle curves	low	10%-20%	strong	medium
3	orange	adequate	low	basic adjustable	limited	sculpted ceiling with gentle curves	normal	<10%	strong	low
4	green	spacious	high	premium adjustable	no	sculpted ceiling with gentle curves	low	<10%	strong	low
5	green	restricted	low	basic adjustable	limited	hard lines or flat surface	relatively low	10%-20%	weak	high
6	green	adequate	high	not adjustable	limited	sculpted ceiling with gentle curves	relatively low	<10%	medium	high
7	pink	spacious	medium	basic adjustable	no	hard lines or flat surface	normal	<10%	medium	high
8	orange	spacious	low	not adjustable	no	sculpted ceiling with gentle curves	relatively low	<10%	weak	medium
9	orange	spacious	medium	not adjustable	rich	sculpted ceiling with gentle curves	low	10%-20%	weak	high
10	blue	restricted	low	not adjustable	no	sculpted ceiling with gentle curves	normal	20%-30%	strong	high
11	pink	adequate	low	premium adjustable	rich	sculpted ceiling with gentle curves	relatively low	20%-30%	weak	low
12	orange	adequate	high	basic adjustable	rich	hard lines or flat surface	low	20%-30%	strong	high
13	blue	spacious	medium	premium adjustable	limited	sculpted ceiling with gentle curves	relatively low	20%-30%	strong	high

Table 6.2 Aircraft cabin interior design configurations for evaluations (Cont'd)

#Configuration	Interior Color	Legroom	Noise	Interior Light	IFE	Interior Pattern	Cabin air pressure	Relative humidity	Vibration	Contaminant
14	Green	Spacious	Low	Premium Adjustable	Rich	Hard Lines or Flat Surface	Normal	10%-20%	Strong	Medium
15	White	Adequate	High	Premium Adjustable	No	Sculpted Ceiling with Gentle Curves	Normal	10%-20%	Weak	High
16	White	Spacious	Low	Basic Adjustable	Limited	Sculpted Ceiling with Gentle Curves	Low	20%-30%	Medium	Medium
17	Pink	Restricted	Low	Premium Adjustable	Rich	Sculpted Ceiling with Gentle Curves	Low	<10%	Medium	High
18	Blue	Adequate	Medium	Premium Adjustable	Limited	Hard Lines or Flat Surface	Low	<10%	Weak	Medium
19	Orange	Restricted	High	Premium Adjustable	No	Hard Lines or Flat Surface	Relatively Low	20%-30%	Medium	Medium
20	Pink	Spacious	High	Not Adjustable	Limited	Hard Lines or Flat Surface	Normal	20%-30%	Weak	Low
21	Blue	Adequate	Low	Not Adjustable	No	Hard Lines or Flat Surface	Low	10%-20%	Medium	Low
22	Pink	Adequate	Medium	Basic Adjustable	No	Sculpted Ceiling with Gentle Curves	Relatively Low	10%-20%	Strong	Medium
23	Blue	Spacious	High	Basic Adjustable	Rich	Sculpted Ceiling with Gentle Curves	Relatively Low	10%-20%	Medium	Low
24	Orange	Restricted	Medium	Premium Adjustable	Limited	Sculpted Ceiling with Gentle Curves	Normal	10%-20%	Medium	Low
25	Blue	Restricted	High	Basic Adjustable	Rich	Sculpted Ceiling with Gentle Curves	Normal	<10%	Weak	Medium
26	White	Restricted	Medium	Not Adjustable	Rich	Hard Lines or Flat Surface	Relatively Low	<10%	Strong	Low
27	Green	Adequate	Medium	Not Adjustable	Rich	Sculpted Ceiling with Gentle Curves	Normal	20%-30%	Medium	Medium

6.3.3 Hypothesis

Based on the discussion above, I summarize four hypotheses addressing the main affective and cognitive influences on choice decision making in the CPT-based preference model:

H1: When the participant is joyful and excited (positive affect), the value function would have a larger α and a smaller β than when the participant is neutral (no particular affect);

H2: When the participant is anxious (negative affect), the value function would have a smaller α and a larger β than when the participant is neutral;

H3: The value function would have a smaller α, a smaller β, and a larger λ, and the weighting function would have a smaller z (δ or θ) of affect-rich products than those of affect-poor products;

H4: The value of λ would be larger than 1, indicating aversion to unpleasant preference.

6.3.4 Emotion Elicitation

Self-elicited methods by imagination or imagery (e.g., Picard et al., 2001; Sinha and Parsons, 1996) are used in this study to elicit target affective states from the participants. It requires participants to be involved in the target affective states deliberately by recalling or imagining a certain situation. In order to facilitate the procedure, two descriptions related to cabin interior design are provided. First, it is assumed that the aircraft cabin designed is employed on a trip to Paris. For joy and excitement, positive descriptions and corresponding images about Paris are provided: *Paris is the world's leading tourism destination. Among Paris' first mass attractions drawing international interests are the Eiffel Tower, the world's most-visited art museum, the Louvre, housing many works of art, including the Mona Lisa and the Venus de Milo statue...* This description projects the expected emotions positively so that participants would have joy, happiness, and excitement about the trip as a result. For anxiety, questions are shown to the participants: *How will I deal with the local language? How expensive will things be? Will I have good weather? Will my bank hold my credit card? Do I forget to pack something?* These questions typically evoke anxiety since they describe situations with unpredictable or uncontrollable events (Seligman, 1975). Finally, neutral is elicited with no description but with a white blank picture. Neutral is used as a baseline level and a control condition for comparison. Furthermore, only those whose self-reported affective states are consistent with the target ones are used in the experiment.

6.3.5 Participants

University students at Georgia Institute of Technology aged between 20 and 30 with gender balance were recruited. Among them, 60 participated in Study 1, and were divided into three groups averagely with gender balance. Participants in each group were asked to elicit one of the target affective states (i.e., joy and excitement, neutral, or anxiety), respectively, and then asked to evaluate different design configurations of the aircraft cabin. In such a way, the group of participants in a neutral state is regarded as a control group and the other two as treatment groups. 40 participated in Study 2, and were divided into two groups averagely with gender balance. Participants in each group were asked to evaluate one kind of aircraft cabin interior designs, i.e., affect-rich or affect-poor, respectively. These two groups are regarded as two different treatments for comparison.

6.3.6 Procedure

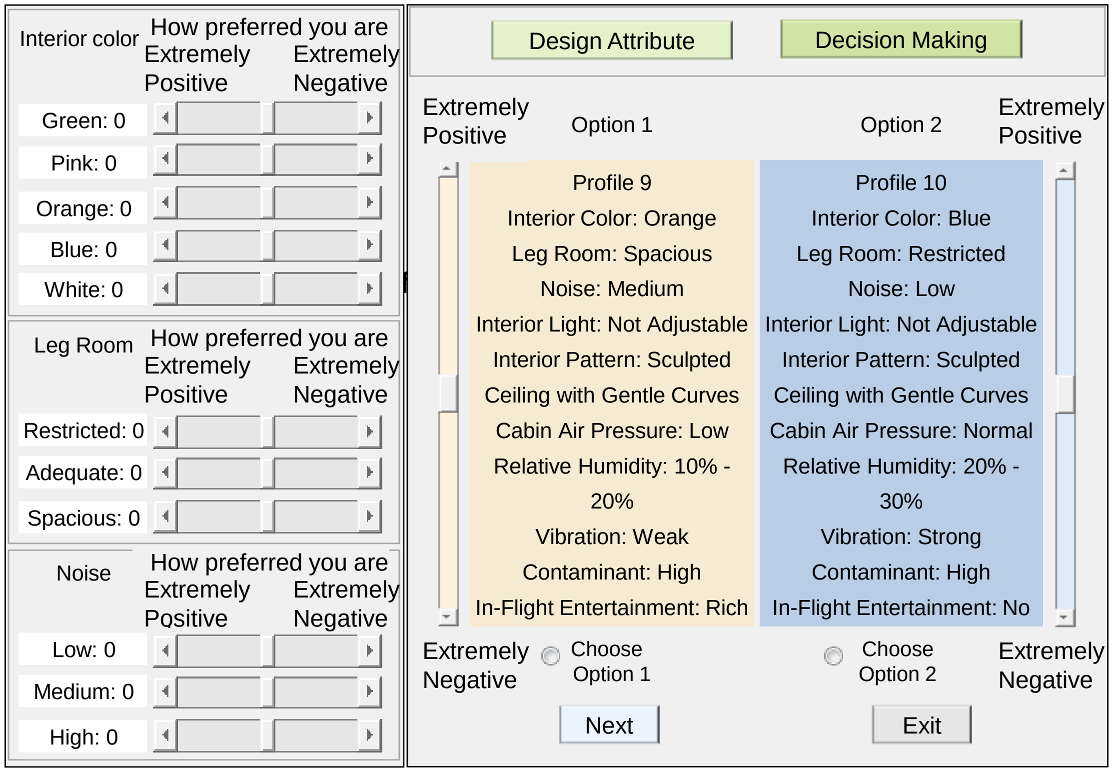
Study 1: After the participants had signed the consent forms, they were told that they were going to take a round- trip flight from Atlanta, USA to Paris, France, and they were asked to self-report cabin preference for different cabin configurations. They were then briefed about the procedure of the study. First, they were asked to navigate a virtual aircraft cabin, focusing on the product attributes. The immersed navigation was operated with a keyboard by the participant who was seated comfortably in an armchair. Although the cabin environment they navigated was not entirely the same as the design configurations offered, it helped them become familiar with the overall environment and improve the accuracy of self-reported preference. Second, they were asked to read the descriptions of the travel (including the associated images) and be involved in the target affective states deliberately by imagining the situation as described. Third, each participant was shown a series of concepts as shown in Table 6.2, and was required to self-report his/her preference level with regard to different product attribute levels and make decisions between two alternative product configurations.

Study 2: All the participants were briefed that they would take a flight for a round trip from Atlanta, USA to Paris, France, and they were required to self-report the cabin preference. However, half of the participants were told that they were on a vacation and beautiful pictures of Paris were shown to them; the rest were told that they were on business with a tight schedule. Other procedures were the same as Study 1 except that they were not asked to elicit emotions. Instead, the first group of participants is considered to self-report the aircraft cabin preference associated with an affect-rich vacation, whereas the second group is associated with an affect-poor trip.

6.3.7 Data Collection

Customer preferences are measured on a scale between -10 and 10, where 10 indicates extremely positive preferences and -10 extremely negative preferences with regard to individual product attribute levels (see Figure 6.5(a)). Further, they are required to make decisions between two alternative design configurations as shown in Figure 6.5(b). Of all the 27 design configurations (in random order for different participants), each participant is required to make 26 decisions (e.g., Configuration 1 vs. Configuration 2, if Configuration 2 is preferred, the next comparison is Configuration 2 vs. Configuration 3, and so on. It thus results in 26 choices). They are preprocessed and structured for analysis of hierarchical Bayesian model to estimate the parameters for different groups of participants.

Based on the experiment of two studies, the data sets for different purposes of the studies are produced. For Study 1, three data sets are generated for three different target emotions, i.e., joyful and excited, anxious, and neutral. For Study 2, two data sets are generated for two types of products, i.e., affect-poor and affect-rich. For both studies, the decision data are divided into a training data set (80% of the data) for parameter estimation and a test data set (20% of the data) for model validation. This process is run three times to generate the averaged results. Two Markov chains are generated. Each chain has 20,000 samples with 5000 burn-in samples, and only every 10th sample is collected. Therefore, for each chain, 1500 samples are valid to estimate the posterior distributions of the parameters. Convergence is confirmed by autocorrelation graphs and visual inspection (Nilsson et al., 2011; Rouder and Lu, 2005).



(a) (b)

Figure 6.5 Data collection: (a) self-reported preference of individual product attributes (only part shown here); (b) Decision making between alternative configurations

6.4 Results and Validation

I present the validation of affective influence on preference-based product choice decision making in terms of whether the results support the hypotheses shown in Section 6.3.3. The validation is also conducted based on the canonical discriminant analysis both for affective states and product types. Finally, the model validation is supported by providing the prediction accuracy in terms of decision making between two alternative product configurations (see Figure 6.5(b)).

6.4.1 Affective Influence

The parameters in the preference model for three different affective states are estimated in Table 6.3, and their posterior density functions are estimated using the *ksdensity* function in Matlab as illustrated in Figure 6.6 (except η). In order to test H1 and H2, ANOVA is used. It shows that there are significant differences among three affective states (α: $F(2,57)$ =19.76, $p < 0.001$; β: $F(2,57) = 6.02$, $p < 0.01$). Bonferroni post-hoc analysis shows that α in anxiety is significantly smaller than those in the other two affective states ($p < 0.05$); β in joy & excitement is significantly smaller than those in neutral and anxiety ($p < 0.01$). The results show that anxious participants tend to be more risk-averse than neutral ones for positive preference, and joyful, excited participants for both positive and negative preference; joyful and excited participants tend to be more risk-seeking than neutral and anxious participants for both positive and negative preference. These results support H1 and H2 (partially). λ values in three affective states are significantly larger than 1, indicating all the participants are averse to negative preference. So H4 is accepted.

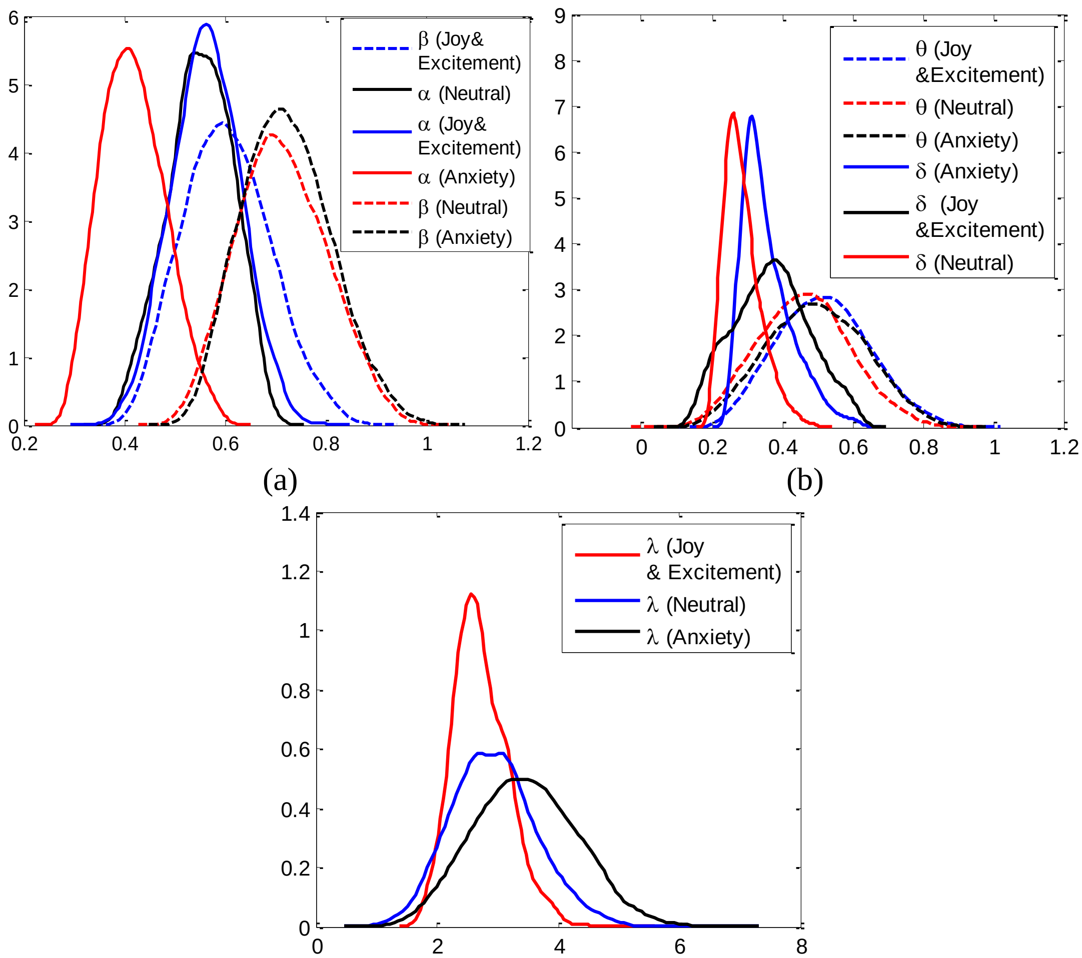
Moreover, for the individual parameters, I conduct a canonical discriminant analysis (Rencher, 1992), in which discriminant functions are linear combinations of individual parameters and are used to predict participants into different affective groups. Figure 6.7 (a) shows the scatter plot of the individual parameters mapped in two canonical discriminant functions. It can be seen that people in joy & excitement and in anxiety seem separated from each other, while people in neutral states seem scattered widely into other two groups. Table 6.4 shows the classification results, in terms of recall, precision, and F-score (see Zhou et al., 2011b). Precision shows how well the model predicts (i.e., a measure of exactness), and recall accounts for how well the model does not miss the target (i.e., a measure of completeness). In Table 6.4, precision and recall are 75.0% and 90.0% for joy & excitement, 66.7% and 60.0% for neutral, and are 88.2% and 80.0% for anxiety, respectively. The F-score combines the precision and the recall and is the harmonic mean of them. It thus gives the optimal accuracy. The mean F-score is 76.3% in Table 6.4.

Table 6.3 Results of parameter estimation in three affective states

Affective states	Parameter mean (standard deviation)					
	α	β	λ	δ	θ	η
Joy & Excitement	0.58 (0.07)	0.62 (0.10)	3.23 (0.83)	0.37 (0.08)	0.53 (0.12)	1.49 (0.85)
Neutral	0.57 (0.07)	0.75 (0.12)	3.08 (0.80)	0.30 (0.06)	0.47 (0.10)	2.17 (0.68)
Anxiety	0.44 (0.06)	0.77 (0.09)	2.86 (0.52)	0.35 (0.07)	0.52 (0.14)	1.65 (0.56)

Table 6.4 Classification based on canonical discriminant analysis for three affective states

	Group	Predicted group membership			# Total instance	Recall
		Joy & Excitement	Neutral	Anxiety		
Original group membership	Joy & Excitement	18	2	0	20	90.0%
	Neutral	6	12	2	20	60.0%
	Anxiety	0	4	16	20	80.0%
	# Predicted instance	24	18	18	60	-
	Precision	75.0%	66.7%	88.2%	-	F-score: 76.3%



(c)

Figure 6.6 Posterior probability density functions for three different affective states: (a) α and β, (b) θ and δ, and (c) λ.

6.4.2 Affect-Rich vs. Affect-Poor Products

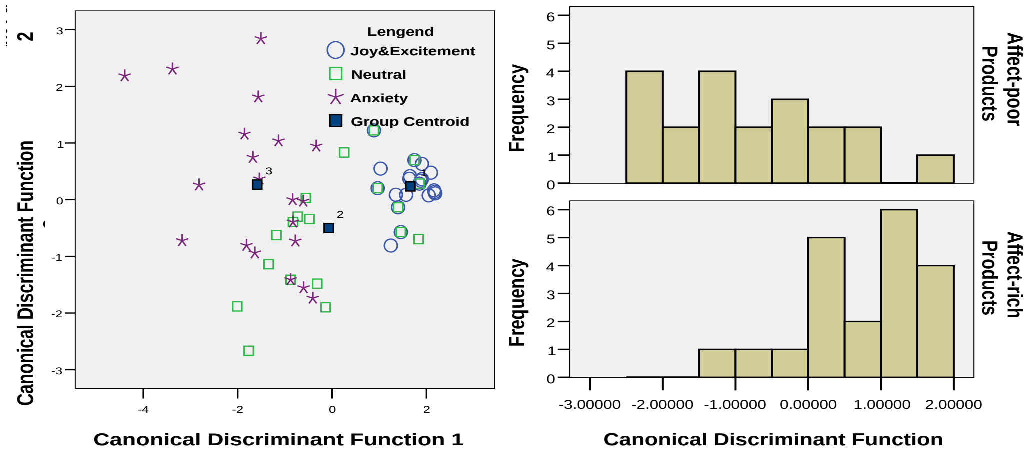
As mentioned previously, two types of products are identified, i.e., affect-poor and affect-rich. The parameters involved in the preference model for these two types are estimated in Table 6.5, and their posterior probability density functions are estimated using *ksdensity* function in Matlab as illustrated in Figure 6.8 (except η). Using ANOVA, it is found that both α and β are significantly smaller (α: $F(1,38) = 9.31$, $p < 0.01$; β: $F(1,38) = 8.26$, $p < 0.01$) in the affect-rich trip than those in the affect-poor trip. This shows that the assumptions of valuation by feeling in affect-rich products and valuation by calculation in affect-poor products are accepted. However, no significant difference is found between the values of λ in two types of products (λ: $F(1,38) = 2.65$, $p = 0.11$). This shows that participants are equally sensitive to two types of products for negative preference. The values of δ and θ are both significantly smaller in the affect-rich trip than those in the affect-poor trip (δ: $F(1,38) = 7.58$ $p < 0.01$; θ: $F(1,38) = 8.65$, $p < 0.01$). This supports (see Figure 6.2) that the weighting function is more curved in the affect-rich trip than that in the affect-poor one. Therefore, H3 is supported except the value of λ. However, λ in two different types of trips are significantly larger than 1, indicating all the participants are averse to negative preferences. Hence, H4 is accepted.

Table 6.5 Results of parameter estimation for two types of products

Product type	Parameter mean (standard deviation)					
	α	β	λ	δ	θ	η
Affect-poor	0.47 (0.06)	0.65 (0.10)	2.79 (0.98)	0.34 (0.09)	0.54 (0.14)	2.11 (0.92)
Affect-rich	0.42 (0.06)	0.51 (0.09)	3.02 (1.01)	0.28 (0.09)	0.44 (0.13)	2.45 (0.83)

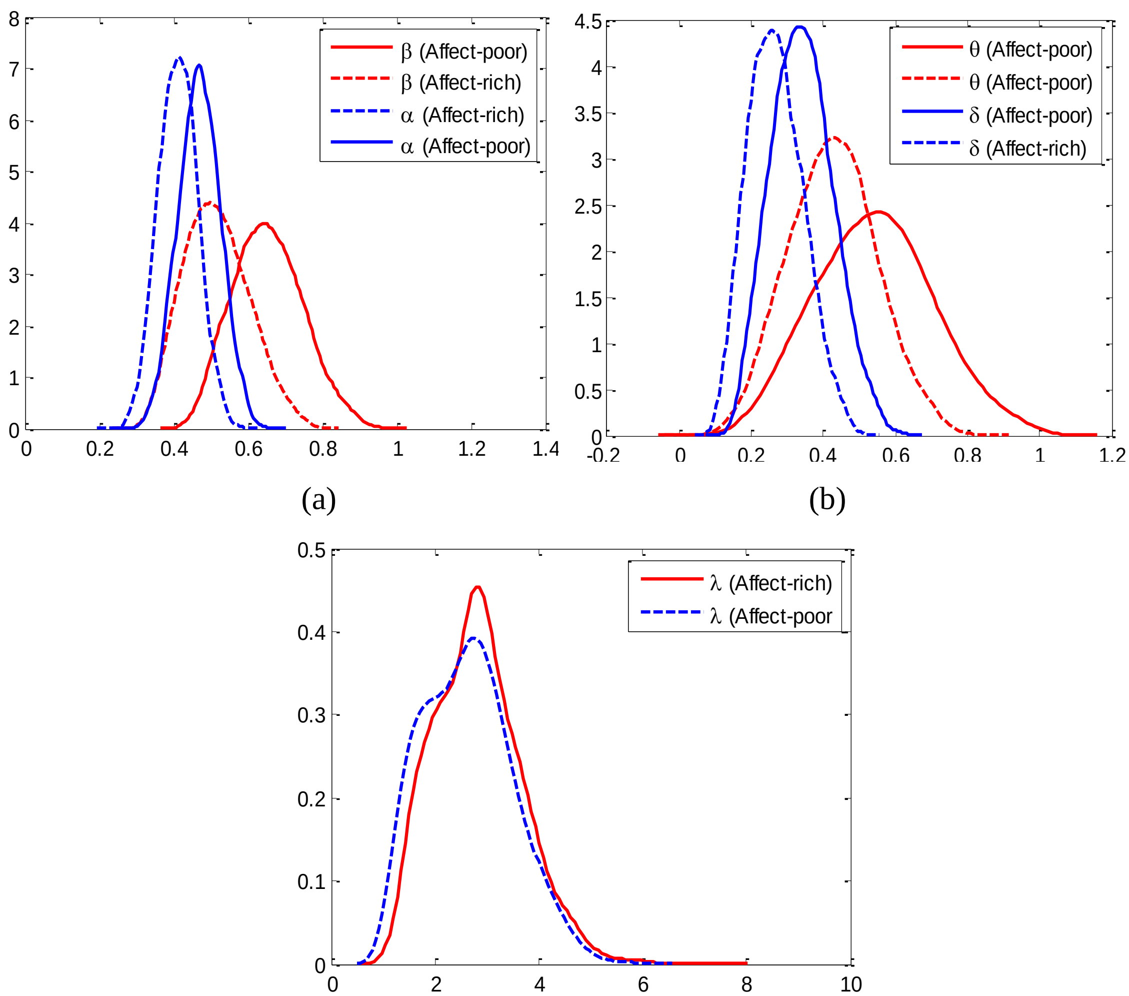
Table 6.6 Classification based on canonical discriminant analysis for two types of products

	Group	Predicted group membership		# Total instance	Recall
		Affect-poor product	Affect-rich product		
Original group membership	Affect-poor Product	17	3	20	85.0%
	Affect-rich Product	6	14	20	70.0%
	# Predicted instance	23	17	40	-
	Precision	73.9%	82.4%	-	$F = 77.4\%$



(a) (b)

Figure 6.7 Canonical discriminant analysis for (a) affective groups and (b) product types



(a) (b)

(c)

Figure 6.8 Posterior probability density functions in two different product types: (a) α and β, (b) θ and δ, and (c) λ

Similarly, I also conduct a canonical discriminant analysis for two types of products (only one canonical discriminant function was produced). Figure 6.7 (b) shows the histogram of the canonical discriminant function in which affect-rich products primarily have positive values, while affect-poor products mainly have negative values. Table 6.6 shows the classification results based on the canonical discriminant analysis. The precision and recall are 73.9% and 85.0% for affect-poor products, and are 82.4% and 70.0% for affect-rich products, respectively. The average F-score is 77.4%.

Besides, I acknowledge the amount of uncertainty associated with each parameter by their posterior distributions as shown in Figure 6.6 and Figure 6.8. Furthermore, substantial uncertainty in drawing inferences about the preference is also indicated by the parameter estimated for individual participants, showing each one has a different preference evaluation function. For example, the mean value of the parameters of participant 8 are α = 0.47, β = 0.76, λ = 2.83, δ = 0.27, θ = 0.24, η = 2.15, and of participant 16 are α = 0.39, β = 0.68, λ = 2.31, δ = 0.30, θ = 0.24, η = 1.98. The values are unique for a particular participant and thus product customization can be realized by applying these parameters for preference evaluation and prediction.

6.4.3 Prediction Accuracy and Optimal Cabin Configuration

In order to validate the proposed model, decision making prediction with a test data set is conducted for three times. The prediction accuracies for different situations are summarized in Table 6.7 for five different models. The prediction accuracy is based on the decision making between the two alternative design configurations (see Figure 6.5(b)) according to Eq. (6.6). It is defined as the number of accurately predicted decisions divided by the total number of decisions made.

Table 6.7 Decision prediction accuracy of different CPT models

Model	Mean prediction accuracy	Standard deviation
Joy & excitement	82.1%	3.2%
Anxiety	83.2%	1.8%
Neutral	77.6%	4.8%
Affect-rich	79.3%	2.9%
Affect-poor	82.2%	2.6%

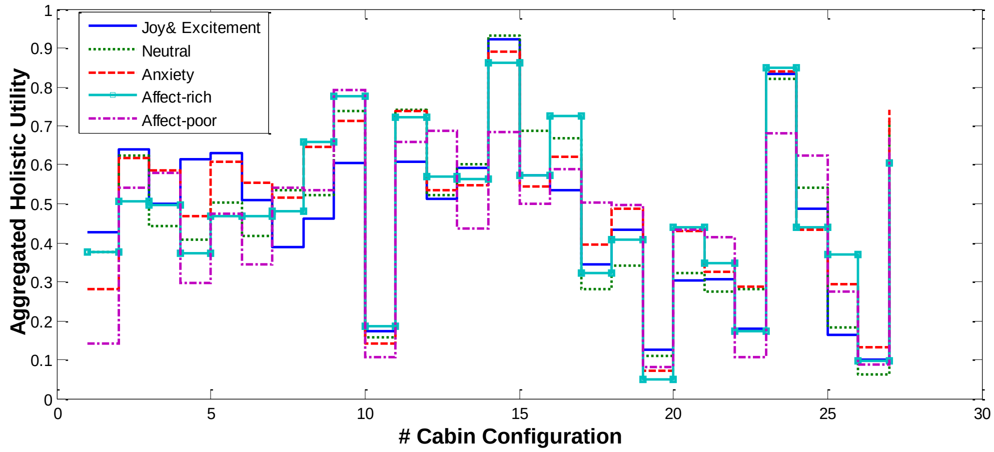


Figure 6.9 Aggregated holistic utility for different cabin configurations of five groups

Furthermore, I obtain the preference prediction functions for all the design configurations in the design space. For example, for the design configurations 1 and 3 in Table 6.2, the evaluations (see Table 6.8) by participants in the neutral group are calculated using Eq. (6.6) with the estimated parameters. The values of aggregated holistic product utilities (indicating customer preferences) for configurations 26 and 27 are 0.10 and 0.65, respectively. Therefore, between the two design options, configuration 27 is preferred. Moreover, for the different participant groups, the aggregated preference (indicating holistic product utility) is shown in Figure 6.9. It shows that the basic trend is consistent among five participant groups and the optimal cabin configurations are configuration 14 (0.89 for anxiety), configuration 14 (0.92 for joy and excitement), configuration 14 (0.93 for neutral), configuration 9 (0.79 affect-poor) and configuration 14 (0.86 affect-rich), respectively.

Table 6.8 Preference comparison between configuration 26 and configuration 27

-	Interior Color	Legroom	Noise	Interior Light	IFE	Interior Pattern	Cabin air pressure	Relative humidity	Vibration	Contaminant
Configuration 26	White	Restricted	Medium	Not Adjustable	Rich	Hard Lines or Flat Surface	Relatively Low	<10%	Strong	Low
Configuration 27	Green	Adequate	Medium	Not Adjustable	Rich	Sculpted Ceiling with Gentle Curves	Normal	20%-30%	Medium	Medium
Reference	0.25	-0.78	-0.58	0.14	-0.24	-0.06	-0.41	-0.12	-0.38	-0.39
Evaluation 26	-0.67	-3.45	-0.48	-1.83	1.67	-1.19	-0.74	-2.15	-2.16	-0.44
Evaluation 27	0.95	0.06	-0.46	-1.81	1.64	1.15	1.00	-0.85	-0.53	1.03

6.5 Discussions

(1) Affective factors in preference-based product choice decision making: Based on the individual parameters, I find that systematical differences exist among the parameters in the subjective value function for participants in three different affective states (see Table 6.3). Consistent with this finding, the canonical discriminant analysis also predicts three affective groups. Specifically, I find that anxious participants tend to be more risk-averse than neutral ones for positive preference, and joyful, excited participants for both positive and negative preference; and that joyful and excited participants tend to be more risk-seeking than neutral and anxious participants for both positive and negative preference. I also identify the differences of six individual parameters with regard to affect-rich and affect-poor products as evidenced in Table 6.5 and Table 6.6. Specifically, participants tend to value affect-rich products by feeling and value affect-poor products by calculation. Furthermore, the results are also validated by decision making prediction based on the parameters estimated. We control the experiment setting strictly the same, except that the participants are in different affective states or product types. In this sense, the differences of the parameters among different groups are mainly attributed to the affective factors rather than cognitive factors.

F-scores for the three affective groups and two types of products are both around 75%, and quite some overlaps of parameter distributions are observed in Figure 6.6 and Figure 6.8. It is possible that the hint of affect elicitation in the experimental setting is so strong that the differences are not that noticeable from the posterior distributions of the parameters. This is also consistent with the results of Figure 6.9. However, in other words, it can be interpreted as the substantial uncertainty involved in the posterior distributions of the parameters. This is otherwise not possible for point estimate in maximum likelihood estimation which often leads to over-confident predictions. As a common practice, the results I obtained are based on the posterior means, which often show the group homogeneity. The hierarchical Bayesian model allows us to calculate parameters for individuals that show individual differences. It inherently implies a personalization strategy for individual customers, which compels the producer to examine different combinations of existing product attributes and value profiles to anticipate and adapt to customers' latent needs. Such a strategy of product innovation is more likely to guarantee positive preferences at both the individual and group levels.

(2) Implications for viral product design: Based on the preference evaluation function, the aggregated perceived preference can be predicted for a particular design configuration, which is a key performance indicator for evaluation of alternative design configurations. Meanwhile, the contribution of individual product attributes to positive or negative preference can also be specified. This provides sensitivity analysis for value-added product design. The results partially support H2 and H3 and support H1 and H4.

There are several implications for viral product design as explained below: First, for three different affective states, the parameters involved in the value function are significantly different. This implies that more investigation needs to be conducted to avoid "affective bias" from customers' subjective perception and recognize the design's actual contribution to customer needs fulfillment in terms of their preferences. Second, affect-rich products usually have a smaller value of α and a larger value of β. It means that the absolute magnitude of positive preferences increases at a slower rate than affect-poor products, and vice versa for negative preferences. However, it will have a larger value of λ, indicating that if one product attribute leads to negative preferences, the holistic utility towards the product will be severely affected. Therefore, more attention should be paid so that no attributes will lead to negative words of mouth in the context of social networks. Third, how the product is framed (i.e., affect-rich or affect-poor) also influences customers' preference perception and evaluation. Therefore, companies should be aware of the fact that products with high affective quality (the ability to elicit positive emotions), but with low probabilities are perceived to have positive preferences, and are thus potentially viral in the social network. For example, for aircraft cabin interior design, first-class cabins with superior product attributes, including spacious personal space, better food and beverage, and booking and check-in priority, etc. are highly deemed in terms of perceived preferences.

(3) Limitations: As an exploratory study, the approach proposed in this research suffers limitations. First, not all the hypotheses are supported. They may be due to the affect elicitation techniques that the flight is hypothetical and the description is plain without vivid multimedia so that the elicited affective states could be different from the target affective states or the intensity is not strong enough. Other possible techniques, such as videos, may be more effective in future studies. Furthermore, the design configurations are mainly based on the word description, which cannot mimic the real aircraft cabin environment and thus lacks ecological validity. In order to improve ecological validity, virtual reality can help create the cabin environment with lower costs in a shorter time compared with real aircraft cabin prototypes.

Second, I only consider three different affective states for a small sample size. It is cautious to generalize the results to other affective states.

Third, there are numerous empirical assumptions made in the hierarchical Bayesian model for parameter estimation and thus the method proposed here may not well accommodate the link between core cognitive decision making and affective influence in terms of parameter shaping. For example, a large amount of subjectivity is involved in selecting prior distributions. In many situations, I believe that proper prior distributions can be prudently selected to effectively make use of the prior knowledge in the estimation process, despite the fact that I can use uninformative priors when no particular prior knowledge is available. In this research, I choose the normal distribution for the group-level parameters between 0 and 1, and the lognormal distribution for the rest based on previous research (Nilsson et al., 2011). I then select either normal or fixed uninformative priors (i.e., uniform distributions) for their priors. These priors actually control the group-level parameters in certain ranges, which provide a very efficient method for estimation since it can fully use the knowledge of the distributional structure (Banks et al., 2005). However, it does not exclude the possibility of better priors. For example, the Beta distribution is often assumed for parameters between 0 and 1 (Bouguila and Elguebaly, 2012).

6.6 Summary

Customer preferences have become the key success factor in product design and adoption and models that can quantify and evaluate preferences have received much attention. This research proposes a CPT-based product choice decision making model for quantifying, predicting, and evaluating customer preferences. The model includes four integrated phases, i.e., the perceptual phase, the affective-cognitive reasoning phase, the learning phase, and the evaluation phase. It addresses the fundamental issues involved in product choice decision making incorporating subjective experiences, including affective and cognitive influences on the CPT-based value function and the weighting function. A hierarchical Bayesian model is utilized to estimate the parameters involved in preference evaluation function, considering the inherent uncertainty by incorporating posterior density functions of the estimated parameters. An empirical study is conducted to test the hypotheses proposed. Results show that affective states influence decision making involved in customer preference-based product choices. The model is also able to predict product choice decision making with moderate accuracies and the optimal product configuration can be identified with the preference evaluation function. I believe that this new model provides the flexibility and comprehensiveness needed to explain decision making behavior involved in customer preference-based product choices.

Therefore, in the context of viral product design, such a model can be used to measure the holistic customers' preference to a product (i.e., holistic utility of a product perceived by a customer). Unlike most diffusion models in viral marketing, which only considers social network effects, a diffusion model that incorporates the holistic customer preference to a product can be more expressive, and thus one's adoption decision can be more informative in the context of online social networks.

CHAPTER 7

A LINEAR THRESHOLD-HURDLE MODEL FOR PRODUCT ADOPTION PREDICTION INCORPORATING SOCIAL NETWORK EFFECTS

With the development of social media, online social networks offer potential opportunities for firms to analyze customer behaviors, especially product adoption, which provides profound technical and economic implications for viral marketing and innovative product design. Among many, one of the fundamental questions is how to predict product adoption, incorporating peer influence of social networks. The answer to this question lays the foundation for product adoption maximization and viral product design in large social networks.

In this chapter, as a way to model social networks, I propose a linear threshold-hurdle (LTH) model to predict product adoption incorporating peer influence of social networks. I attack multiple limitations of traditional diffusion models by modeling activation thresholds, influence probability, adoption spread, and holistic utility of the product in a holistic fashion. First, I propose finer activation thresholds based on the five categories of adopters. In addition, I identify three operational factors underlying social network effects, including interaction strength, structural equivalence, and social entity similarity, to model influence probabilities. Furthermore, I distinguish influence spread from adoption spread by introducing a tattle state, in which customers express opinions without adopting the product. Finally, I introduce the notion of hurdle to capture the monetary aspect in customers' decision making process of product adoption. Based on the proposed LTH model, two data mining methods based on the rough set technique, namely, decision rules and decomposition trees, are employed to predict product adoption in a large social network. An empirical study of Kindle Fire HD 7 inch tablets is used to illustrate the potential and feasibility of the proposed model. The results demonstrate the predictive power of the proposed model with an average F-score of 89.8% for the week prediction model and 86.7% for the bi-week prediction model.

7.1 Peer Influence on Product Adoption

A modern product like an iPhone or iPad works not only because of its inherent industrial and interface design, but also because of the social networks in which it "lives" (Cho et al., 2010). With the pervasive connectivity of the Internet and social media, including review sections of online shopping websites (e.g., Amazon.com) and online social networks (e.g., Facebook), customers become more interconnected and informed when they make product choices. The social network plays a fundamental role as a medium for the spread of information, ideas, and influence among its social entities (Kempe et al., 2003). In this process, the social entities consider not only the attributes of a product, but also the preferences and influence of other customers in the social network. As mentioned earlier in Chapter 1, there are three effects identified in this process, including word-of-mouth effects, imitation effects, and network effects (Dou et al., 2013), and they can be understood as social network effects. They often take place when customers aspire to be like or unlike others, or learn something new about certain products from others. There is little doubt that one can hardly isolate his purchase or usage decisions from his social networks. Such effects often lead to the spread of adoption behavior from one social entity to another in the social network (Kleinberg, 2008).

The increasing availability of social network data has drawn more attention to understand the social network effects on customers' product adoption decisions and adoption maximization (Iyengar et al., 2011). One of the fundamental questions is how to predict product adoption for social entities who have not adopted by now (Fang et al., 2013). The answer to this question is not only critical to viral marketing and design with regard to product adoption in social networks (e.g., Aral and Walker, 2011; Chen et al., 2010), but also vital to applications in demand estimation (e.g., Hartmann, 2010), public health (e.g., Chen et al., 2011b) and politics (e.g., Bello and Rolfe, 2014), etc. For example, in viral marketing, it is important to identify a set of powerful influencers in the social network as seeds so that the expected number of social entities who adopt the product can be maximized. This is dependent on the reliable prediction of product adoption in the search process of optimal seeds, because one needs to predict how likely other entities will adopt if the initially targeted ones adopt (Fang et al., 2013).

Given the limitations of the current diffusion models (both independent cascade models and linear threshold (LT) models) described in Chapter 4, including influence probability modeling in an ad-hoc fashion, inability to distinguish between influence spread and adoption spread, and inability to incorporate negative reviews and comments that discourage others to adopt a product, I propose a LTH model to overcome these limitations, and a data mining method named rough set is used to predict product adoption in a large social network with a case study of Kindle Fire HD 7 inch tablets. The social network is constructed based on the reviewer-commenter information from Amazon.com. Our empirical study offers two interesting results. First, the proposed models can have the best F-scores of 89.8% and 86.7%, when the prediction time windows are within one week and two weeks, respectively. These results are significantly better than models that do not incorporate the features related to the notion of hurdle. Second, the results predicted within one week are better than those predicted within two weeks. These findings suggest that it is possible to predict product adoption by leveraging the information within in a large social network, i.e., peer influence, based on the proposed LTH model.

7.2 Problem Formulation and System Architecture

A social network is denoted with a directed graph $G = (V, E)$ with a set of N social entities, $V = \{v_1, v_2, \dots, v_N\}$, and a set of edges E, showing the social ties among social entities. The social network is dynamically evolving with regard to time in a discrete fashion, $t = 0, 1, 2 \cdots$, during which social entities adopt a product successively. The adoption prediction problem can be formulated as follows:

Given the information at the current time T in the social network, including

1) The adoption decision and adoption time of each social entity up to T,

2) The social network structure up to T,

3) The comments and reviews of each social entity about the product up to T, and

4) The characteristics associated with each social entity up to T.

Predict the adoption decision at time $T + 1$ for those who have not adopted up to T.

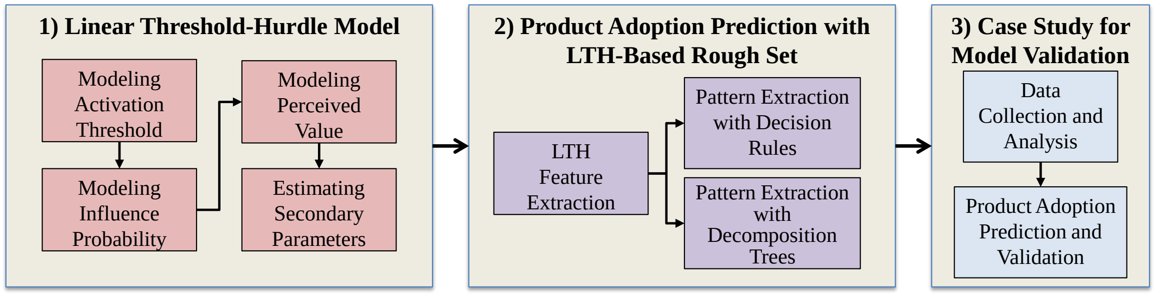


Figure 7.1 Overview of the system architecture

Based on the problem formulation, the system architecture to solve the problem is illustrated in Figure 7.1. It consists of three modules. The first module is the proposed model, i.e., LTH model. It has four components, including modeling an activation threshold of an inactive social entity, modeling the influence probability from the active neighbors to the inactive social entity, modeling holistic utility of the product, and estimating the secondary parameters. These four components are linked in a holistic fashion and are more expressive than the original LT model. Therefore, the proposed model attempts to overcome the limitations mentioned above. The second module is product adoption prediction with LTH-based rough set. It has three components. The first component makes use of the features from the LTH model, which will be used for pattern extraction in the following components. The last two components are parallel with two rough set-based data mining methods, including decision rules and decomposition trees. The third module is the case study for model validation. First, data are collected with regard to a case study of Kindle Fire HD 7 inch tablets, and then data are analyzed with regard to the first two modules. Using the two data mining methods based on the rough set technique, prediction results are obtained and two kinds of validation plans are conducted.

7.3 Linear Threshold-Hurdle Model

Based on previous studies (Bhagat et al., 2012; Fang et al., 2013; Lu and Lakshmanan, 2012; Rogers, 2003), Figure 7.2 shows the overview of the proposed LTH model in terms of a state diagram and its influence factors. Let $G = (V, E, L)$ denote a weighted, directed graph with a set of vertices V (indicating social entities) and a set of edges E (indicating social ties in social networks); the influence probability is represented by a function $L: E \rightarrow [0,1]$. Each vertex v_i has an activation threshold θ_i, and when it is smaller than the influence probability p_i from the social network, it will become active. Then the customer v_i has a holistic utility U_{ij} of the product P_j (indicating customer preference) and a hurdle utility h_{ij} of the product P_j. Here the hurdle utility is understood as the highest price that a customer is willing to pay for the holistic utility of the product. Since holistic utility of a product is relative to a customer's hurdle utility. I assume that the holistic utility of a product is influenced by the reviews and comments from his social neighbors, whereas a customer's hurdle utility is fixed and is equal to the price of the product (see Figure 7.2). Then, when the holistic utility is larger than the price Δ_j of the product P_j, i.e., $U_{ij} > \Delta_j$, the customer will adopt the product; otherwise the customer will enter a state called *tattle*, in which the customer will promote the product with some probability μ_{ij}, otherwise the customer will demote the product, i.e., giving negative product reviews and/or comments. In the tattle state, the customer comments based on the information about the product probably heard from other people, without purchasing the product or physically interacting with the product. However, when the customer adopts the product, based on his personal user experience, he or she will promote the product with some probability ρ_{ij}, otherwise the customer will demote the product, i.e., giving negative reviews or comments. These reviews and comments then further influence the holistic utility of other customers in the social network.

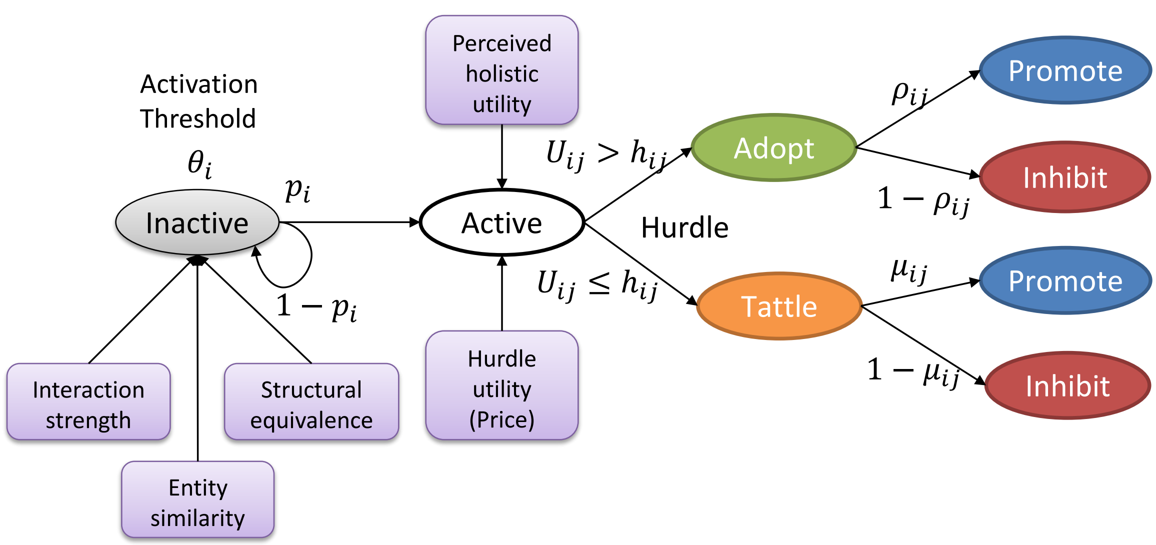


Figure 7.2 Overview of the linear threshold-hurdle model

7.3.1 Activation Threshold

As mentioned in Chapter 3, Rogers (2003) suggests five categories of adopters, including innovators, early adopters, early majority, late majority, and laggards. Each category of adopters has different activation thresholds. Based on the *S*-curve of the adoption process (see Figure 7.3), with successive groups of customers adopting the product (shown in blue), its market share (yellow) will eventually reach the saturation level (Fisher and Pry, 1971). We define the activation thresholds for the successive five categories mentioned above as following the distributions: Uniform (0, 0.025), Uniform (0.025, 0.16), Uniform (0.16, 0.50), Uniform (0.50, 0.84), Uniform (0.84, 1), respectively. A data mining method named rough set (Pawlak, 1991) is used to identify the adopter category for each customer based on the historical activities performed within the online social network. The detailed description of rough set is shown in Section 7.4.

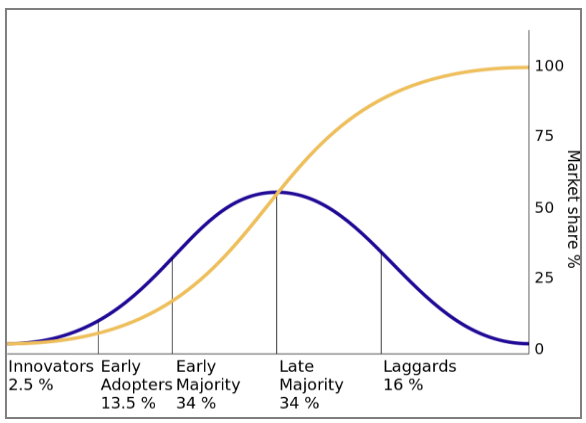


Figure 7.3 The diffusion of innovations, according to Rogers (2003)

7.3.2 Influence Probability

Based on the social influence network theory (Friedkin, 1998) and (Fang et al., 2013), I identify three operational factors to model influence probabilities, including social interaction strength, social network structural equivalence, and social entity similarity.

Peer influence within a social network is firstly reflected by the social ties resulting from investing time and reciprocity (Levy, 1992). It makes sense that people connected by stronger ties have greater influences on one another than those connected by weaker ties (Wellman, 1997). The social tie between two social entities can be directional (e.g., Twitter) or non-directional (e.g., Facebook). Let $V = \{v_1, v_2, \dots, v_N\}$ be a set of social entities in a social network, and x_{ji}^t denote the strength of the social tie from v_j to v_i at time t. Note for directional ties, x_{ij}^t is often different from x_{ji}^t. We propose to measure x_{ji}^t by the number of interactions between two entities at time t, which is normalized as the following (Fang et al., 2013):

$$I_{ji}^t = (x_{ji}^t - x_{min})/(x_{max} - x_{min}), \tag{7.1}$$

where x_{max} and x_{min} represent the maximum and the minimum social tie strength, respectively. Hence, the larger the value of I_{ji}^t, the stronger the social interaction between v_i and v_j.

Moreover, the opinions and behaviors between social entities are also affected by the structure of a social network itself. Burt (1987) suggests that structurally equivalent entities in the social network are likely to make similar judgments, even without direct communication, because these two social entities connect to other entities identically. In other words, when they are in the same position in the social structure, they are prone to the same influence from other entities, or even mimic others with which they interact (Wasserman and Faust, 1994). According to Burt (1987), structural equivalence between social entities can be measured by Euclidean distance d_{ji}^t between v_i and v_j at time t. For a social network with directional social ties, d_{ji}^t is:

$$d_{ji}^t = \sqrt{\sum_{v_k \in V/\{v_i, v_j\}} [\left(c_{ik}^t - c_{jk}^t\right)^2 + \left(c_{ki}^t - c_{kj}^t\right)^2]}, \tag{7.2}$$

and for a social network with non-directional social ties , d_{ji}^t is computed as:

$$d_{ji}^t = \sqrt{\sum_{v_k \in V/\{v_i, v_j\}} \left(c_{ik}^t - c_{jk}^t\right)^2}\ , \tag{7.3}$$

where $c_{ab}^t = 1$ if there is a social tie from v_a to v_b at time t; and $c_{ab}^t = 0$, otherwise. The larger the value of d_{ji}^t, the less structural equivalence between v_i and v_j. To be consistent with I_{ji}^t, d_{ji}^t is normalized by:

$$E_{ji}^t = (d_{max} - d_{ji}^t)/(d_{max} - d_{min}), \tag{7.4}$$

where d_{max} and d_{min} are the maximum and minimum distances among the entities in the social networks, respectively.

Furthermore, homophily also has been proved to be a cause for similar behaviors (McPherson et al., 2001). Despite the fact that social influence and homophily are different (Aral et al., 2009), both of them can be useful in predicting customers' future behavior. Specifically, similar choice behavior may result from similar opinions, interests, and attitudes towards products and services (Aral et al., 2009). In this research, I focus on predicting product adoption without necessarily distinguishing influence from homophily. Social entity similarity is measured by the distance between two social entities (Hand et al., 2001). Assuming two social entities are associated with m_e characteristics (e.g., age, location, interests, degree, and the like), i.e., $v_i = \left(c_1^i, c_2^i, \dots, c_{m_e}^i\right)$ and $v_j = \left(c_1^j, c_2^j, \dots, c_{m_e}^j\right)$. If the *i*-th characteristic is nominal, then $S\left(c_i^i, c_i^j\right) = 0$ when $c_i^i = c_i^j$, and 1 otherwise. If it assumes a real value or integer number, then $S\left(c_i^i, c_i^j\right) = |c_i^i - c_i^j|/(max_i - min_i)$, where max_i and min_i are the respective maximum and minimum values of the *i*-th characteristic among all the entities. Then the similarity between social entities v_i and v_j at time t can be denoted as:

$$S_{ji}^t = 1 - \sum_{k=1}^{m_e} S\left(c_k^i, c_k^j\right)/m_e. \tag{7.5}$$

It shows that the larger the value of S_{ji}^t, the more similar between v_i and v_j.

Based on these three operational factors, I compute the influence probability from social entity v_j to social entity v_i at time t as a logistic function, i.e.,

$$p_{ji}^t = \frac{1}{1+\exp\left[6-4\left(I_{ji}^t+E_{ji}^t+S_{ji}^t\right)\right]}. \tag{7.6}$$

Note that the expression $6-4\left(I_{ji}^t + E_{ji}^t + S_{ji}^t\right)$ is in such a form, so that it is within $[-6, 6]$. Hence, the Eq. (7.6) approximates p_{ji}^t between 0 and 1, and when $I_{ji}^t + E_{ji}^t + S_{ji}^t$ closes to 3, p_{ji}^t tends to be 1. Then the influence probability from v_i's active neighbors $N_i^{a,t}$ to v_i at time t, i.e., p_i^t can be calculated as

$$p_i^t = 1 - \prod_{v_j \in N_i^{a,t}} \left(1 - p_{ji}^t\right). \tag{7.7}$$

7.3.3 Holistic Utility

The social entity v_i's holistic utility of the product P_j at time t, i.e., U_{ij}^t, is influenced by the reviews and comments from v_i's active neighbors (including social entities who adopt and tattle about the product) in the social network. These reviews and comments can be positive or negative. Usually, ratings are provided by social entities when they review the product. I assume whenever ratings are missing, they are predicted using sentiment analysis of the reviews and comments (Zhou and Jiao, 2014). The higher the ratings are, the more positive the reviews and comments are. For a Likert-scale system, assuming the highest rating is R_h and the median rating is R_m (indicating neutral opinion), the holistic utility of v_i is calculated as follows:

$$U_{ij}^{t} = \Delta_j \frac{R_h - R_m + \bar{R}_{v_k \in N_i^{a,t}}}{R_h}, \tag{7.8}$$

where Δ_j is the price of the product P_j and $\bar{R}_{v_k \in N_i^{a,t}}$ is the mean rating by v_i's active neighbors at time t. Intuitively, if $\bar{R}_{v_k \in N_i^{a,t}} < R_m$, the holistic utility of v_i will be smaller than the price of the product, and vice versa.

7.3.4 Secondary Parameter Estimation

Secondary parameters in the LTH model include the probabilities of promoting product P_j when v_i is in the states of adopt and tattle, respectively, i.e., ρ_{ij} and μ_{ij}. These probabilities are empirically defined as

$$\rho_{ij} = \frac{\#\text{positive review of adopted products}}{\#\text{review of adopted product}}, \tag{7.9a}$$

$$\mu_{ij} = \frac{\#\text{positive review of nonadopted products}}{\#\text{review of nonadopted product}}. \tag{7.9b}$$

Despite the fact that these parameters are not used in the model of adoption prediction directly, they influence the holistic utility of the product. In this sense, they influence the prediction implicitly.

7.4 Product Adoption Prediction with LTH-Based Rough Set

7.4.1 Data Feature Extraction from LTH Model

Based on the proposed LTH model, eight features are extracted for each social entity v_i, within his or her active neighbors $N_i^{a,T}$ at current time T, including interaction strength I_i^T, entity similarity S_i^T, structural equivalence E_i^T, activation threshold θ_i, influence probability p_i^T, whether $\theta_i < p_i^T$, $\bar{R}_{v_j \in N_i^{a,T}}$ the mean rating, and whether $\bar{R}_{v_j \in N_i^{a,T}} < R_m$. Note the first three features are the joint influence or similarity of $N_i^{a,T}$

with regard to v_i, and they are defined as $I_i^T = \sum_{v_j \in N_i^{a,T}} I_{ji}^T$, $S_i^T = \sum_{v_j \in N_i^{a,T}} S_{ji}^T$, and $E_i^T = \sum_{v_j \in N_i^{a,T}} E_{ji}^T$.

7.4.2 LTH-Based Rough Set for Adoption Prediction

Two data mining methods based on the rough set (Pawlak, 1991) technique are adopted to construct the prediction models, including decision rules and decomposition trees. Assume an information system, $I^S = (\Pi, F)$ such that $\forall \mathbf{f} \in F: \Pi \to F^*$, where Π is a non-empty finite set called the universe, F is a non-empty feature set, for any feature vector $\mathbf{f} = \{f_l\}_L \in F$, where $L = 8$ is the total number of the features involved, and F^* is the value set of F. Corresponding to the feature vector, a decision vector $\mathbf{d} = \{d_h\}_H \in D$ with H decision variables is defined, where $H = 1$ is the total number of the decision involved, i.e., adoption decision. For the rough set theory, its input is tabulated as a decision table, $\Gamma = (F \cup D, I)$, where $F \cup D$ is the universe of inference I. For the i-th training sample in Γ, $I_i^* \sim (F_i^* \to D_c^*)$, embodies an inference relationship from the features F_i^* to the corresponding decision D_c^*. $F_i^* = \{f_{il}^*\}_L$ is the value set of the feature vector $\mathbf{f}$; $D_c^* = \{d_c^*\}_C$ is the value set of $\mathbf{d}$, and $D_c^* = \{\text{adopter}, \text{nonadopter}\}$ in this study.

(1) Decision rule: IF-THEN rules are used to predict product adoption. Rule generation is based on the concept of reduct that is a subset of attributes in a decision table that can fully characterize the knowledge contained in the decision table. A reduct is defined as a subset of features in I^S, such that $\Phi = \{\phi_k\}_K \subset \Gamma$, where $\phi_k = (f_i^\phi, d_i^\phi)$ is subject to an indiscernibility relation, in which, for objects $x \in \Pi$ and $y \in \Pi$, a pair $(x, y) \in \Pi \times \Pi$ belongs to Φ. Therefore, for any object $z \in \Pi$, a decision rule can be generated, such that $\forall q \in [1, Q]$, where Q denotes the total number of features instantiated by this rule, the predecessor of the rule takes the conjunction of certain feature instances, $f_q^\phi(z)$, and the successor takes on specific values of decision variables, $d^\phi(z)$. The general form of a decision rule constructed for reduct Φ and object z is thus given as below:

$$IF\left(f_1^\phi = f_1^\phi(z)\right) \& \left(f_2^\phi = f_2^\phi(z)\right) \& \cdots \& \left(f_Q^\phi = f_Q^\phi(z)\right) THEN(d^\phi = d^\phi(z)). \quad (7.10)$$

It is often useful to discretize the features with cuts into different intervals if they are numeric and the rough set theory does it by using the global strategy based on the maximal discernibility heuristics (Pawlak, 1991). These rules provide bases to predict product adoption. For a given test sample, *tst*, the subset of rules matched by *tst* is selected. If *tst* matches only rules with the same adoption decision, then the one predicted by those rules is assigned to *tst*. However, if *tst* matches rules with different adoption decisions, a commonly used measure in Eq. (7.11) for conflict resolution is to be made so that the adoption decision with the highest measure value is chosen (Gora and Wojna, 2002).

$$Strength(tst, c) = |\bigcup_{r \in MatchRules(tst, d_c^*)} SupportSet(r)|, \quad (7.11)$$

where d_c^* denotes the *c*-th adoption decision, *SupportSet*(r) is a set of training examples matching the rule, r, *MatchRules*(*tst*, d_c^*) is a subset of minimal rules that are applicable to *tst* and the adoption decision is c, and $|\cdot|$ denotes the cardinality of a set '·'.

The training process is to extract all the IF-THEN rules from the training set. When $t = 0$, the training set is empty. When $t = 1, 2, 3 \cdots$, I calculate the feature vectors for all the social entities $v_i \in V_N^t$, where V_N^t is the set of non-adopters up to t in the form of $[I_i^{t-1}, E_i^{t-1}, S_i^{t-1}, \theta_i, p_i^{t-1}, \theta_i < p_i^{t-1}, \bar{R}_{v_j \in N_i^{a,t-1}}, \bar{R}_{v_j \in N_i^{a,t-1}} < R_m \rightarrow$ adoption decision at $t]$. All the feature vectors form the decision table at time T, i.e., Γ_T. This decision table is then used to extract all the IF-THEN rules described above.

The testing process can be summarized as follows:

1) Given the training set obtained at T;
2) Feature vector $[I_i^{t-1}, E_i^{t-1}, S_i^{t-1}, \theta_i, p_i^{t-1}, \theta_i < p_i^{t-1}, \bar{R}_{v_j \in N_i^{a,t-1}}, \bar{R}_{v_j \in N_i^{a,t-1}} < R_m]$ is calculated for those social entities $v_i \in V_N^T$, who have not adopted up to T;
3) Predict $v_i \in V_N^T$ would adopt the product or not at $T + 1$ using the IF-THEN rules extracted.

(2) Decomposition tree: Decomposition is to partition a large data table into smaller ones with common features (Bazan and Szczuka, 2001). The structure is similar to a binary decision tree in which each route from the root to the leaf leads to the final class, i.e., adoption or not. However, unlike decision trees (e.g., C4.5, ID3), at the leaf stage, the final class is predicted by the decision rules produced by the rough set theory. The tree's every internal node is labeled by some template (any non-leaf vertex of the tree), and every external node (leaf) is associated with a set of training samples matching all the templates in a path from the root to a given leaf. The decomposition process is described in Algorithm 1 and the prediction process is described in Algorithm 2 (Bazan and Szczuka, 2001).

Algorithm 1 Decomposition by template tree:

Step 1: Find the best template tem in the decision table Γ;

Step 2: Divide Γ into two subtables (or sub-trees) containing all training samples matching and not matching template tem, respectively, i.e., $\Gamma(tem)$ and $\Gamma(\overline{tem})$;

Step 3: If the subtables are of a self-defined size, then generate decision rules for $\Gamma(tem)$ and $\Gamma(\overline{tem})$, respectively, else $\Gamma = \Gamma(tem)$ and $\Gamma = \Gamma(\overline{tem})$, and repeat Step 1 to Step 3, respectively.

Algorithm 2 Prediction by decomposition tree:

Step 1: If for any given test sample, tst, matches tem in Γ, then goes to $\Gamma(tem)$, else goes to $\Gamma(\overline{tem})$;

Step 2: If tst is at the leaf of the tree, then goes to step 3, else $\Gamma = \Gamma(tem)$ or $\Gamma = \Gamma(\overline{tem})$, and repeat Step 1 and Step 2;

Step 3: Classify tst using the decision rules attached to the leaf.

7.5 Case Study

A case study of Kindle Fire HD 7 inch tablets (released on September 14, 2012) is used to illustrate the proposed method. In order to construct the social network regarding this product, I make use of the reviewer-commenter information about this product from Amazon.com. Figure 7.4 shows a typical review of the product and two comments are associated with it. Thus, two links from this reviewer are connected to two commenters. We also see this reviewer is a verified purchaser, showing that this is an adopter. This information can be used as ground truth and allows us to validate the model.

49 of 59 people found the following review helpful

☆☆☆☆☆ **Love it!**, October 13, 2013

By **Amazon Customer "Jamieann"** - See all my reviews

Verified Purchase (What's this?)

This review is from: Kindle Fire HD 7", HD Display, Wi-Fi, 8 GB - Includes Special Offers (Electronics)

I actually purchased this as a gift for my daughter's 6th birthday and she is in love with it! She is just starting out with reading on her own and this gave her some extra encouragement. We haven't had any issues with it so far. The speed seems great to me. We haven't done much web browsing though, so I can't give any feedback on that. She has just been using the Kindle Free Time to access a ton of free educational games, videos, and books and it has been great for that! This is my first Kindle and so far I am very happy with it. It is serving the purpose I bought it for and for $139 I think it is a great deal!

Help other customers find the most helpful reviews

Was this review helpful to you? Yes No

Report abuse | Permalink

Comments (2)

Figure 7.4 A typical review about the Kindle Fire HD 7 inch tablet

7.5.1 Data Collection

I collect the reviewer-commenter information from September 2012 to September 2013 (50 weeks) using Python 2.7 (www.python.org). Figure 7.5 shows the overall social network, which has 5220 vertices and 10476 edges. The figure is generated using NodeXL with the Fruchterman-Reingold layout (Fruchterman and Reingold, 1991) in the form of different groups obtained with the Clauset-Newman-Moore clustering algorithm (Clauset et al., 2004). Figure 7.6 shows the degree distribution (here I consider it as a non-directional graph) on a log-log scale. It resembles the familiar power law distribution. Besides, it has a giant connected component along with a large number (i.e., 244) of smaller connected components. Hence the social network generated is consistent with past studies on large-scale social networks (Leskovec and Horvitz, 2008). I further plot the product awareness (in terms of commenting on Kindle Fire HD tablets) and adoption process from September 2012 to September 2013 as illustrated in Figure 7.7. The awareness process is decreasing over time while the adoption process mimics Figure 7.3 except at the end. Close examination upon this product, it is possibly due to the fact that the next generation, i.e., Kindle Fire HDx, was released in September 2013 while the price of the product in this study was reduced substantially, which boosted the sales to a large extent. Figure 7.8 shows the histogram of days from product awareness to product adoption. The maximum value is 311 days, the minimum value is 0 day, the median value is 140 days, and the mean value is 124.64 days. This information also gives hints to identify adopter categories.

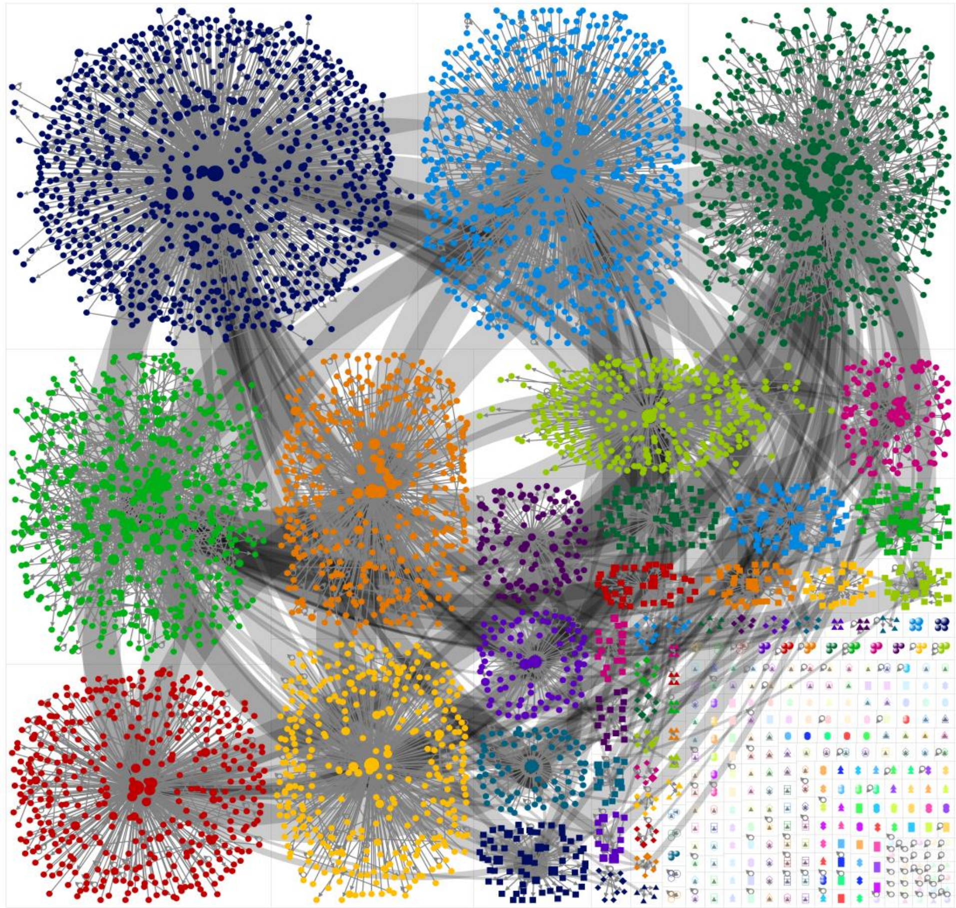


Figure 7.5 The constructed social network based on the reviewer-commenter links about the Kindle Fire HD 7 inch tablet from Amazon.com from September 2012 to September 2013

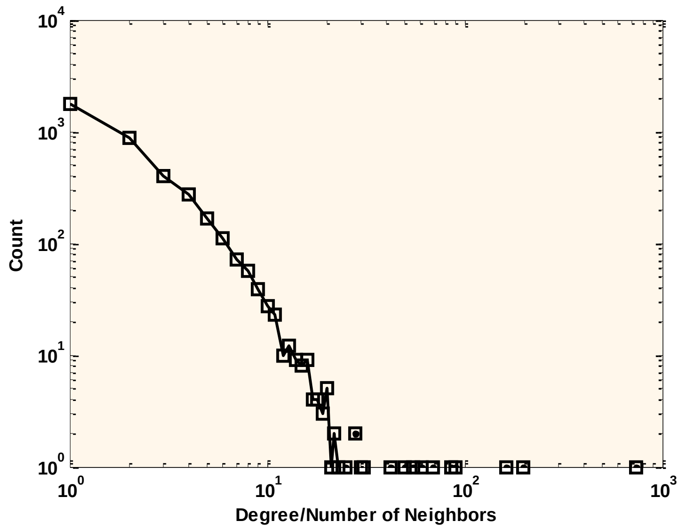


Figure 7.6 Degree distribution on a log-log scale

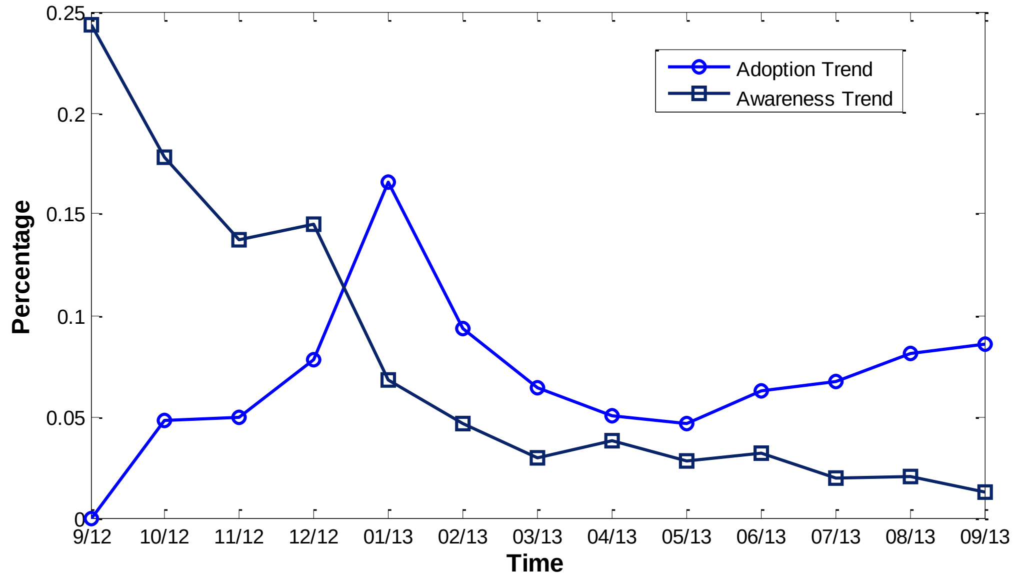


Figure 7.7 The awareness and adoption process from September 2012 to September 2013

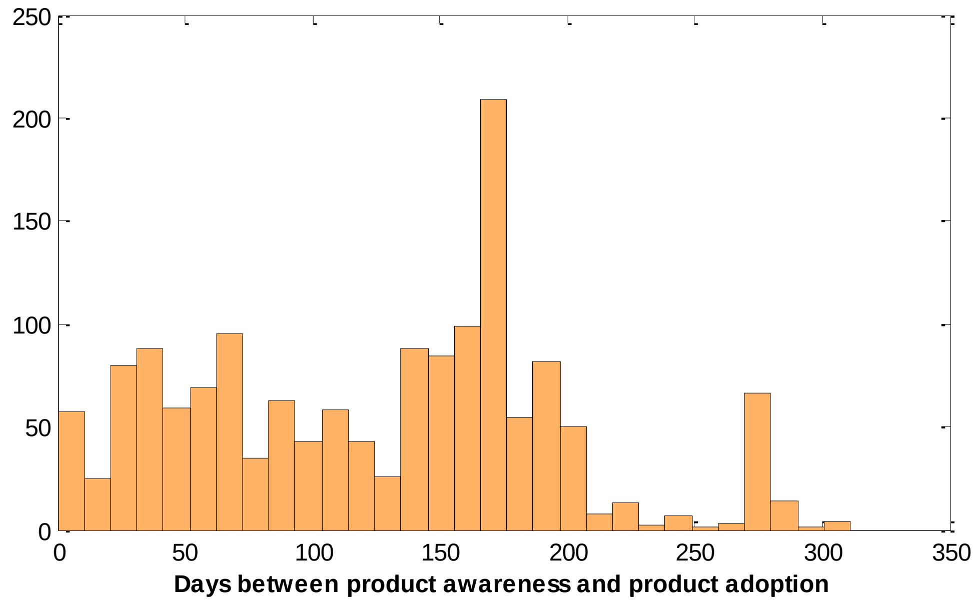


Figure 7.8 Histogram of days from product awareness to product adoption

7.5.2 Data Analysis

In order to determine which adopter category a social entity belongs to, I capitalize on the information associated with each customer crawled from Amazon.com using Python, including geographic information, interests, number of reviews, top reviewer ranking, number of comments, helpfulness of the reviews, and total votes received in terms of how helpful the review is. Then, decision rules based on the rough set technique are used to predict the adopter category of each social entity so as to determine the activation threshold distribution. The influence probabilities are obtained based on the method described in Section 7.3.2. Note for entity similarity, I make use of both graphic measures for each social entity and entity characteristics. The former consists of three centrality measures, i.e., closeness, betweenness, and eigenvector, and a cluster coefficient (Newman, 2010); the latter includes those used to determine which adopter category a social entity belongs to as mentioned above. From Amazon.com, each review gives a certain number of stars using a 5-point Likert scale, where 1 star is least satisfied and 5 stars most satisfied. Hence, $R_h = 5, R_m = 3$. I can calculate U_{ij}^t/Δ_j, rather than U_{ij}^t in Eq. (7.8). Finally, ρ_{ij} and μ_{ij} can be obtained from the profile of each social entity from Amazon.com.

7.6 Results and Model Validation

7.6.1 Validation Plan

In this study, two major validation plans are carried out. The first one is to test whether the proposed model, i.e., the LTH model is better than the original LT model. Therefore, two types of feature vectors are employed. The proposed model in this research includes all the eight features as described in Section 7.4.1, whereas the original LT model excludes two features related to the notion of hurdle, i.e., $\bar{R}_{v_j \in N_i^{a,T}}$ the mean rating, and whether $\bar{R}_{v_j \in N_i^{a,T}} < R_m$. The second one is to test the sensitivity of time in product adoption prediction. Two types of prediction models are constructed, including a week prediction model and a bi-week prediction model. Based on the data collected, I decompose them into 50 weeks, representing 50 social networks constructed at the end of the week. For the week prediction model, I pick one week as the current time T ($T = 1, 2, \ldots, 50$), and the data by the end of week T as the training data. Then the two data mining methods based on the rough set techniques, i.e., decision rules and decomposition trees, are used to predict product adoption in week $T + 1$ for those who have not adopted by the end of week T. Likewise, for the bi-week prediction model, the data are divided into 25 bi-weeks, representing 25 social networks constructed at the end of the bi-week. I pick two weeks as the current time T ($T = 1, 2, \ldots, 25$), and the data by the end of bi-week T as the training data. The same data mining methods are used to predict product adoption in bi-week $T+1$ for those who have not adopted by the end of bi-week T.

I use F-score to measure the accuracy of the prediction results. F-score is defined as the harmonic mean of precision and recall, i.e., F-score = $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$ (Zhou et al., 2014d). Precision is a measure of exactness or fidelity, whereas recall is a measure of completeness.

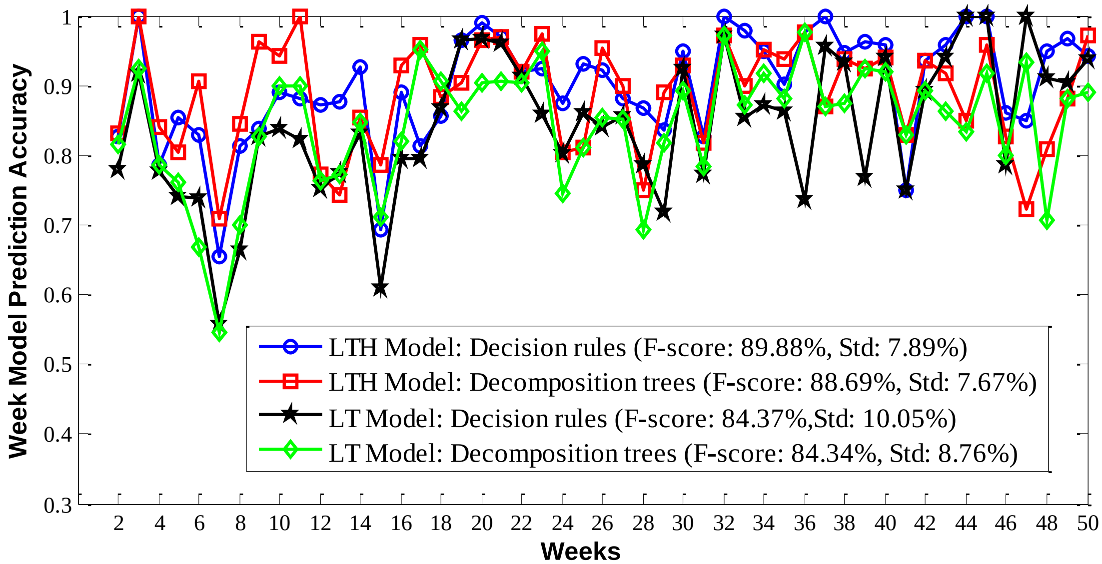
7.6.2 Results for Validation

(1) LT model vs. LTH model: Figure 7.9 (a) shows the comparison between the LTH model and the LT model in terms of the F-scores for the week model. As for the LTH model, I can see that the mean F-scores for the decision rules and the decomposition trees are 89.88% and 88.69%, respectively. Compared with those F-scores for the LT model (84.37% and 84.34%, respectively), using t test for two independent samples (they are considered as two independent samples as they are obtained by two different models), I find that the decision rule method of the proposed model significantly outperforms those of the original model ($t(74) = 3.39, p < 0.01$); similarly, the decomposition tree method of the proposed model performs significantly better than those of the original model ($t(74) = 2.63, p < 0.05$). Therefore, for the week prediction model, the proposed LTH model significantly outperforms the original LT model.

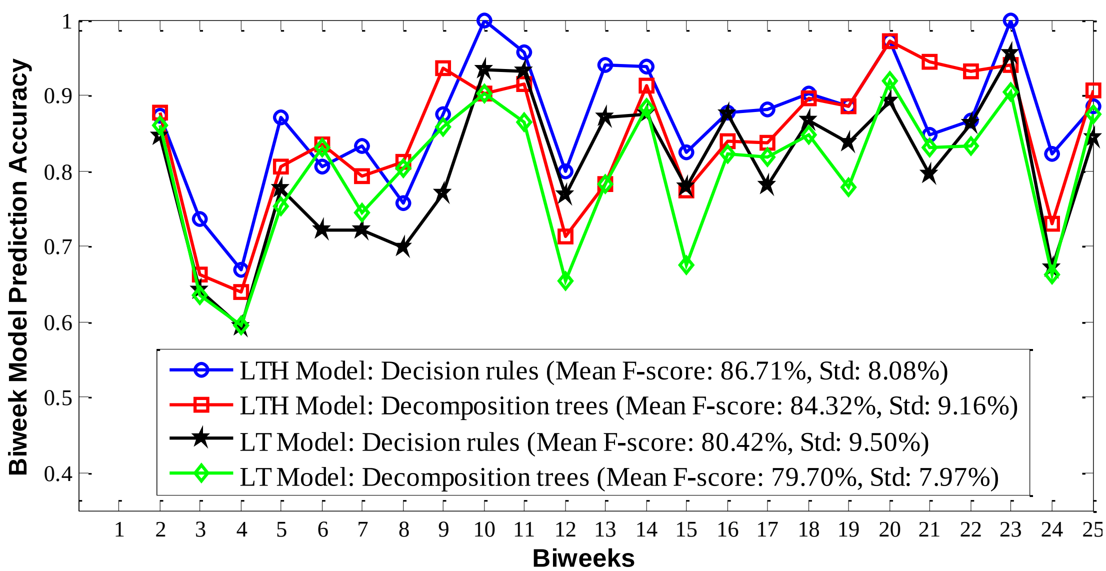
Figure 7.9 (b) shows the comparison between the LTH model and the LT model in terms of the F-scores of the bi-week model. As for the LTH model, I can see that the mean F-scores for the decision rules and the decomposition trees are 86.71% and 84.32%, respectively. Compared with those F-scores for the LT model (80.42% and 79.70%, respectively), using t test, I find that the decision rule method of the proposed model significantly outperforms those of the original model ($t(46) = 2.47, p < 0.05$); the decomposition tree method of the proposed model performs marginally better than those of the original model ($t(46) = 1.74, p < 0.10$). Therefore, for the bi-week prediction model, the proposed LTH model at least marginally outperforms the original LT model.

(2) Week prediction model vs. Bi-week prediction model: Furthermore, I compare the F-scores between the week prediction model and the bi-week prediction model, as demonstrated in Table 7. 1. The models are indicated with different colors in

Figure 7.9. From the comparison results using *t* test, I can tell that the week prediction model at least marginally performs better than bi-week prediction model. This seems to be plausible since it would be easier to predict product adoption within a week than to predict product adoption within two weeks.



(a)



(b)

Figure 7.9 Comparison between LTH model and LT model in terms of *F*-scores. (a) The week prediction model; (b) The bi-week prediction model

Table 7.1 Comparison between week prediction model and bi-week prediction model

Week vs. Bi-week	Blue (89.88%) vs. Blue (86.71)	Red (88.69%) vs. Red (84.32%)	Black (84.37%) vs. Black (80.42%)	Green (84.34%) vs. Green (79.70%)
$t(70)$	1.67	2.19	1.68	2.09
p value	< 0.10	< 0.05	< 0.10	< 0.05

7.7 Discussions

(1) LTH model: In order to deal with the limitations discussed of the traditional diffusion models, this chapter proposes an LTH model to predict product adoption within a large social network based on rough set theory. First, I model the activation threshold based on the adopter categories proposed in the communication and innovation diffusion studies (Rogers, 2003). This manipulation allows us to model the activation threshold of each customer in a finer fashion compared with the traditional uniform distribution between 0 and 1. Second, three operational factors are identified based on established social network theories, including interaction strength, structural equivalence, and entity similarity. Among them, social features, network measures, and customer characteristics are all included. These operational factors are then used to model the influence probabilities. Therefore, this method is more expressive than traditional influence probability modeling methods. Third, the proposed model distinguishes between influence spread and adoption spread by incorporating both the tattle state and the adoption state. In such a way, not all the active social entities go to the adoption state automatically and unconditionally. Fourth, I propose the notion of hurdle by making use of the comparison between customers' holistic utility and the price of the product. This notion captures the monetary aspect and the product aspect in the decision making process of product adoption.

Based on the empirical study of Kindle Fire HD 7 inch tablets from Amazon.com, I am able to validate the proposed model. Two rough set-based data mining methods, namely, decision rules and decomposition trees, demonstrate the predictive power of the proposed method. Specifically, the proposed LTH model performs significantly better than the original threshold model for the week model and marginally better for the bi-week model. Interesting findings are also obtained that the week model at least marginally outperforms the bi-week model.

(2) Implications for viral marketing: Our method can give insights into how the marketer can incorporate peer influence of social networks into product marketing decisions for viral marketing and demand estimation. An example of such phenomenon is the mobile phone family plans that encourage using the same phone network by offering large discounts to customers for in-network calls (Gunnec, 2012). For another example, when a member of a social group has to make a choice among the products, it is natural to take into account the positive or negative reviews and comments (information sharing among customers over social media (e.g., Facebook) or review sections of shopping websites (e.g., Amazon) (Liu et al., 2013)), in addition to considering the product's functional attributes.

In terms of viral marketing, the product adoption prediction problem is the fundamental one in maximizing product adoption within a social network. By understanding the product adoption intention of each social entity, one can effectively identify the seeds that are used to influence other customers in the social network to adopt a product. Furthermore, different incentive strategies can be deployed for individual customers based on their product adoption intension in order to maximize sales and minimize marketing costs. On the other hand, to better understand the demand estimation, the fundamental question is still how to predict future adoption for individual customers who have not adopted by now. By answering this question, different resources and capabilities can be allocated dynamically and across different time frames based on the prediction results.

(3) Limitations: The proposed method also suffers several limitations. First, more factors can be incorporated in the prediction model to be more expressive. For example, it is reported that the strength of weak ties (Granovetter, 1973) and the content analysis of product reviews and comments (Zhou and Jiao, 2014) influence product adoption in the context of social networks.

Second, it is also possible that not all the adopters will review or comment on the product after adopting a product (Romero et al., 2011). In order to deal with this problem, I exclude all the adopters who did not express their opinions in the study. However, we might split the adopt state into two sub-states, namely, "adopt and review" and "adopt and not review" (Bhagat et al., 2012). Such a model tends to be more expressive and consistent with the real situation.

Third, the ratings provided by tattlers may not be as trustworthy as those of adopters, since they do not interact with the product physically. In this sense, these people are often biased with their opinions, such as loyal fans of Apple products. Therefore, more investigation is needed in this aspect. On the other hand, I can make use of the content of reviews and comments offered by the adopters and the tattlers rather than simple ratings to estimate customers' holistic utility of the product. Furthermore, content analysis can help understand the trustworthiness and helpfulness of the opinions provided by the tattlers (Liu et al., 2013).

Fourth, the proposed model needs to be validated with many more data sets from different types of social networks with different types of products. Rogers (2003) show five intrinsic characteristics of innovations that influence an individual's decision to adopt or reject an innovation, such as complexity. Therefore, a tablet and a music CD are two different types of products in terms of their complexity. Furthermore, different types of social networks may have different impacts on product adoption prediction, such as random, scale-free, and small word networks. Besides, our social networks can make use of ratings and customer opinions from Amazon.com while it would be difficult for other social networks, such as Facebook.

Finally, the proposed method needs scalable heuristics algorithms to handle very large networks. For example, in order to calculate structural equivalence, the distance calculation needs $O(N^2)$, where N is the total number of vertices in the social network. One possible method is to restrict the size when calculating structural equivalence for approximation.

7.8 Summary

With the proliferation of social network analysis for studying customer behaviors in large social networks, it is increasingly important to understand how customers spread the influence and the adoption of the product with expressive models. This chapter proposes a LTH model incorporating peer influence to predict product adoption in the context of a large social network. We understand the product adoption process in two stages. The first stage mainly triggers awareness of the product that creates influence spread. In the second stage, customers compare their holistic utility with the price of the product so as to decide whether to adopt the product or not. The user experiences and opinions with the product with or without adopting the product further influence other customers' holistic utility of the product. Based on the case study, it is rather promising for companies to leverage social network effects in predicting product adoption and then propose strategic plans for viral marketing and design in the context of large social networks.

CHAPTER 8

BI-LEVEL GAME THEORETIC OPTIMIZATION FOR VIRAL PRODUCT DESIGN EVALUATION

Viral product design involves sophisticated interactions between product portfolio planning and viral marketing. However, social network effects are mainly considered in marketing related activities, and there is still limited investigation of the interplay between product design and viral marketing. In the context of social networks, it is important to leverage both viral product attributes and viral influence attributes for social network effects so that both product adoption and product line performance can be optimized jointly. In order to deal with the joint optimization problem, this chapter presents a systematic formulation of a Stackelberg game theoretic optimization model for viral product design evaluation. The product adoption maximization problem is modeled as the leader and the product portfolio optimization problem is modeled as a follower. The interaction and coupling of these two optimization problems are addressed with a coordinate-wise optimization strategy iteration by iteration, in which adoption maximization is tackled with an improved greedy algorithm (i.e., CELF++) and a hybrid Taguchi genetic algorithm (HTGA). A case study of Kindle Fire HD tablets demonstrates the feasibility and potential of bi-level game theoretic optimization for viral product design evaluation, which is advantageous over the existing viral marketing that only considers viral influence attributes, i.e., seed customers.

8.1 Marketing-Engineering Coordination for Viral Product Design

Social network research has been increasingly attracting the attention of scholars from the domains of marketing, sociology, and engineering, and so on (Aral et al., 2013). One of the good reasons is that social networks have revolutionized the way how humans interact, such as that between companies and customers and among customers themselves. In particular, one of the important problems is influence maximization (InfMax) in viral marketing (Bhagat et al., 2012). However, as mentioned previously, one of the limitations of InfMax is that previous researchers only consider social network effects based on viral influence attributes in the product diffusion and adoption process, without taking into account customers' preferences to products, especially viral product attributes that can spread widely in a social network. The share-of-choice (SoC) problem is another important research task in product line design, in which multiple product variants are optimized in order to maximize the number of product adopters (Kohli and Krishnamurti, 1989). The SoC problem shares the same objective with the InfMax problem. Despite the fact that the SoC problem considers the customer preferences in terms of product attributes, it ignores the interactive relations between customers. Given these two questions, it is plausible that they can be combined due to the fact that they are mutually complementary. Therefore, it is important to combine these two important aspects for viral product design, which is referred to as adoption maximization with viral attributes (AdpMaxVA).

However, the SoC problem involves combinatorial optimization of attribute levels for product configurations. These attributes are subject to engineering constraints and goals, such as customer satisfaction and costs. These engineering concerns must be taken into consideration in the design and manufacturing process in order to maximize product line performance. In this sense, viral product design involves sophisticated marketing-engineering coordination. The marketing goal aims to maximize product adoption, whereas the engineering goal aims to maximize product line performance in terms of costs and customer satisfaction. Therefore, it is imperative to coordinate marketing and engineering concerns and to optimize these two problems jointly that bridge the gaps between these two domains. Towards this end, in this chapter,

(1) I formulate viral product design as a bi-level game theoretic optimization problem for the purpose of evaluation. Both the upper-level and lower-level objectives will be the evaluation criteria. The objective of viral product design is to select an optimal set of viral product attributes and attribute levels for product portfolio planning so as to maximize product adoption. In addition to a standard product portfolio planning problem, viral product design entails another adoption maximization problem that aims to find an optimal set of viral product attributes and viral influence attributes (i.e., seed customers) for maximizing the expected number of product adopters. The interplay of product design

and viral marketing necessitates an integration of these two optimization problems within a coherent framework of joint optimization.

(2) We propose a solution strategy to tackle the bi-level optimization problem. In order to solve this type of complex optimization problems efficiently, I adopt a coordinate-wise optimization strategy (Barbieri and Bonchi, 2014) to fix one type of variables (i.e., seed set or product attribute) in the optimization problem. When the product attributes are fixed, it becomes an InfMax problem, which is solved by the CELF++ algorithm (Goyal et al., 2011). When the seed set is fixed, the bi-level optimization problem is first transformed into a single-level parametric optimization problem, which is solved by a HTGA (Tsai et al., 2004).

(3) I conduct a thorough case study with a real-world semantically rich data set from Amazon.com to demonstrate the potential and feasibility of the proposed method. The social network is constructed based on the reviewer-commenter links about Kindle Fire HD tablets, and the candidate sets of product attributes and attribute levels are derived based on the review content. The customer profiles on Amazon.com, operational factors of social network effects, and graphic metrics help facilitate data analysis in the case study.

8.2 Leader-Follower Joint Optimization

Viral product design involves a leader-follower joint optimization problem, that is, to optimize product portfolio planning and maximize product adoption jointly. It leverages both marketing and engineering concerns as shown in Figure 8. 1. In this research, the leader F leverages both viral product attributes and viral influence attributes for product adoption maximization as the marketing goal. The basic idea is to identify an optimal set of attributes and attribute levels and an optimal set of seed customers that can maximize product adoption in the social network. Thus, the leader F has an s-dimensional design variable $\mathbf{X} \in \mathbf{R}^s$. The follower f has a t-dimensional design variable $\mathbf{Z} \in \mathbf{R}^t$, including attribute levels and product variants**.** The follower considers both customer satisfaction and engineering cost that can maximize the expected overall performance of the entire product line as the engineering goal. Between the leader and the follower, the social network is where the decision is tested. A leader-follower model assumes certain decision power for both the leader and the follower, with the leader possessing a higher priority.

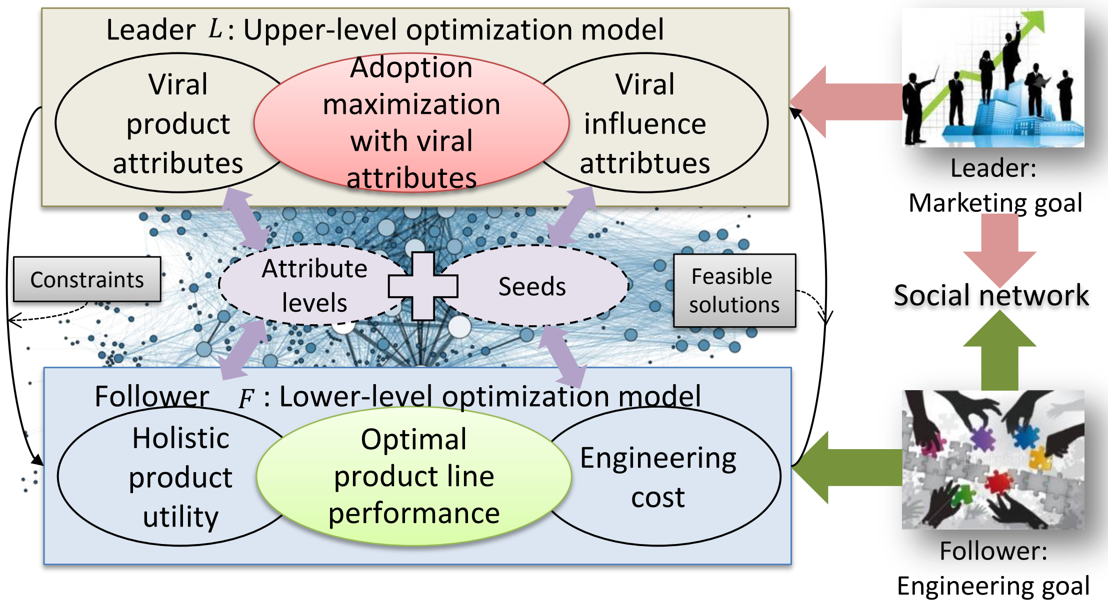


Figure 8.1 System model of bi-level decision making for viral product design

Generally speaking, the bi-level formulation can be represented as follows (Bard, 1998; Colson et al., 2007):

$$\mathrm{Max}_{\mathbf{X}}\ F(\mathbf{X},\mathbf{Z}), \tag{8.1a}$$

$$\text{s.t.}\quad G(\mathbf{X},\mathbf{Z}) \leq 0, \mathbf{X} \in \mathbf{R}^{s}, \tag{8.1b}$$

where $\mathbf{Z}$ solves

$$\mathrm{Max}_{\mathbf{Z}}\ f(\mathbf{X},\mathbf{Z}), \tag{8.1c}$$

$$\text{s.t.}\quad g(\mathbf{X},\mathbf{Z}) \leq 0, \mathbf{Z} \in \mathbf{R}^{t}, \tag{8.1d}$$

where G and g are vector valued functions of dimension p and q, showing the constraints. From Eq. (8.1) I can denote the constraint region as $\Omega = \{(\mathbf{X},\mathbf{Z}): G(\mathbf{X},\mathbf{Z}) \leq 0, g(\mathbf{X},\mathbf{Z}) \leq 0, \mathbf{X} \in \mathbf{R}^{s}, \mathbf{Z} \in \mathbf{R}^{t}\}$. The projection of Ω onto the upper-level design space gives the feasible set for $\mathbf{X}$, i.e., $\mathrm{U} = \{\mathbf{X} \in \mathbf{R}^{s}: \exists \mathbf{Z} \in \mathbf{R}^{t}, \text{such that } (\mathbf{X},\mathbf{Z}) \in \Omega\}$. Then the follower's rational reaction set for $\mathbf{X} \in \mathrm{U}$ can be defined as $\mathrm{R}(\mathbf{X}) = \{\mathbf{Z} \in \mathbf{R}^{\boldsymbol{t}}: \mathbf{Z} \in \mathrm{argmin}\{f(\mathbf{X},\bar{\mathbf{Z}}): g(\mathbf{X},\bar{\mathbf{Z}}) \leq \mathbf{0}\}\}$. The most important concept is the inducible region, which defines the feasible set at least when the lower-level optimization model has a unique optimal solution for all values of $\mathbf{X}$, which can be defined as

$$\mathrm{IR} = \{(\mathbf{X},\mathbf{Z}): (\mathbf{X},\mathbf{Z}) \in \Omega, \mathbf{Z} \in \mathrm{R}(\mathbf{X})\}. \tag{8.2}$$

I assume that the bi-level optimization problem is well-posed in that Ω is nonempty bounded, and IR is nonempty. In addition, I assume that $\mathrm{R}(\mathbf{X})$ is single-valued, which implies that there exists a unique response function, $\mathbf{Z} = \mathbf{Z}'(\mathbf{X})$**.** Then, I can find an optimal solution to Eq. (8.1), denoted as is $(\mathbf{X}^{*}, \mathbf{Z}^{*})$ based on the joint optimization of product adoption and product portfolio planning.

Solving the bi-level model can be realized in three steps:

Step 1: The leader makes a decision, $\mathbf{X}'$, according to the leader's own strategy $F(\mathbf{X}, \mathbf{Z})$; and then announces the decision to the follower with a set of constraints;

Step 2: The follower makes a decision subject to his/her own strategy f together with the leader's decision; and then feedback the follower's feasible solution, $\mathbf{Z} = \mathbf{Z}(\mathbf{X}')$, to the leader; and

Step 3: The leader adjusts its decision to obtain a new $\mathbf{X}''$, based on the follower's feasible solution.

These steps are iterated until a satisfactory result is arrived for both the leader and the follower (Ji et al., 2013). However, such solution is often not efficient. Based on the unique response function $\mathbf{Z} = \mathbf{Z}'(\mathbf{X})$**,** Eq. (8.1) can be converted into a single-level parametric optimization problem (Colson et al., 2007) as follows:

$$\text{Max } F(\mathbf{X}, \mathbf{Z}'(\mathbf{X})), \tag{8.3a}$$

$$\text{s.t.} \quad G(\mathbf{X}, \mathbf{Z}'(\mathbf{X})) \leq 0, \mathbf{X} \in \mathbf{R}^{s}. \tag{8.3b}$$

Despite the fact Eq. (8.3) is essentially a bi-level optimization problem, it paves the way for solving such problems without going through the three steps directly, and algorithms, such as evolutionary algorithms are proposed to solve the single-level parametric optimization problems efficiently (Li et al., 2014).

8.3 Game Theoretic Optimization for Viral Product Design

8.3.1 Upper-Level Optimization Model

(1) The SoC problem: In order to formulate the upper-level optimization model, I repeat the SoC problem and the AdpMaxVA problem. Viral product design involves selecting an optimal set of viral product attributes and attribute levels. This is related to product portfolio planning, which aims at the selection of a near-optimal mix of product variants configured by different product attribute levels to offer in the target market (Jiao et al., 2007b). A product line involves multiple product variants (e.g., 16GB, 32GB, and 64GB models of Kindle Fire HD tablets), which attract heterogeneous customers in the market. Each product variant is configured with K product attributes, i.e., $a_k \in A, k = 1, \ldots, K$. Let L_k be the number of levels for the k-th attribute. Then the attribute level set can be represented as $A^* = \{a^*_{kl} | k = 1, \ldots, K, l = 1, \ldots, L_k\}$. Considering all the possible configurations, there are a number of meaningful J product configurations, indicated by $X = \{X_j | j = 1, \ldots, J\}$, and $X_j = (X_{jk} | k = 1, \ldots, K) = (x_{j11}, \ldots, x_{j1L_1}, \ldots, x_{jK1}, \ldots, x_{jKL_K})$ is a vector showing a particular configuration for the j-th product P_j, where $x_{jkl} | k = 1, \ldots, K, l = 1, \ldots, L_k$ is 1 if the l-th attribute level of the k-th attribute is selected for P_j, otherwise it is 0. $Y = (y_j | j = 1, \ldots, J)$ is a vector indicating a particular choice of P_j, and $y_j = 1$ if P_j is chosen, otherwise it is 0.

In the literature of product portfolio planning, the basic principle of identifying an optimal set of product attributes and attribute levels is the SoC problem, aiming to maximize the number of customers who will adopt the product (Kohli and Krishnamurti, 1989). For a set of customers, $V = \{v_1, v_2, \ldots, v_N\}$, they have different part-worth utilities, i.e., u_{ikl}, which denotes the utility of the l-th attribute level of k-th attribute perceived by the i-th customer in the social network. The SoC adoption model postulates that a customer v_i adopts the j-th product only if his/her holistic utility of the product U_{ij} exceeds his hurdle utility, h_{ij}. In this situation, $z_{ij} = 1$, otherwise it is 0. The objective of the SoC problem is to maximize the number of product adopters, i.e., $\sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}$, denoted as $\sigma(X)$. Therefore, the SoC problem can be formulated as follows (Du et al., 2014; Gunnec, 2012):

$$\text{Max} \quad \sigma(X) = \sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}, \tag{8.3a}$$

$$\text{s.t.} \quad U_{ij} = \sum_{k=1}^{K} \sum_{l=1}^{L_k} u_{ikl}\, x_{jkl} \geq h_{ij} z_{ij}, \forall i \in [1, N], j \in [1, J], \tag{8.3b}$$

$$\sum_{l=1}^{L_k} x_{jkl} = 1\,, k \in [1, K], j \in [1, J], \tag{8.3c}$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L_k} \left| x_{jkl} - x_{j'kl} \right| > 0\,, j \neq j', \tag{8.3d}$$

$$\sum_{j=1}^{J} y_j \leq J^+ < J, \tag{8.3e}$$

$$z_i = \sum_{j=1}^{J} z_{ij}\,, j \in [1, J], \tag{8.3f}$$

$$z_{ij}, \in \{0,1\}, x_{jkl} \in \{0,1\}, \forall i \in [1, N], j \in [1, J], k \in [1, K], l \in [1, L_k]. \quad (8.3g)$$

The exclusive constraint (8.3c) guarantees that only one attribute level is selected for each product attribute. Constraint (8.3d) denotes that at least one attribute level is different for two different product variants. J^+ in constraint (8.3e) is the upper bound for the total number of product variants offered in the market.

(2) The AdpMaxVA problem: Viral product design also necessitates an AdpMaxVA problem, that is, how to select a set of key customers to act as the seeds to influence the largest number of customers so that they are likely to become product adopters due to the joint influence of viral product attributes and viral influence attributes. Given a directed social network $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_N\}$ is a set of nodes, indicating a set of customers and a link from v_j to v_i, i.e., $(v_j, v_i) \in E$ means that v_j can potentially influence v_i. Given a certain budget, the key is to find a seed set $S \subseteq V$ with $|S| = n < N$, so that by activating them, I can maximize the expected number (denoted as $\sigma(S)$) of customers that eventually get activated based on a diffusion model. Assume I_i^t is the peer influence of v_i from his or her active neighbors at time t. I_i^t can be calculated by aggregating individual influence from active neighbor v_j to v_i, i.e., I_{ji}^t at time t. I introduce the adoption probability by integrating InfMax and SoC (Barbieri and Bonchi, 2014):

$$\Pr\left(P_j \middle| i, t\right) = \frac{\exp\left(\varphi_i^t(P_j)\right)}{1+\exp\left(\varphi_i^t(P_j)\right)}, \quad (8.4)$$

where

$$\varphi_i^t\left(P_j\right) = I_i^t + U_{ij} - h_{ij}, \quad (8.5)$$

where $\Pr\left(P_j \middle| i, t\right)$ is the probability that v_i will adopt product P_j at time $t + 1$ if it is no smaller than the adoption threshold ϑ_i at time t. In this case, $z_{ij} = 1$ if the product is P_j, otherwise it is 0. Here $\varphi_i^t\left(P_j\right)$ consists of peer influence at time t, i.e., I_i^t, customer i's preference to product j, i.e., U_{ij} (see Section 8.3.2), and customer i's hurdle utility of product j, i.e., h_{ij}. This formulation enables both viral product attributes and viral influence attributes to be incorporated in the product adoption model, and thus referred to as adoption maximization with viral attributes, i.e., AdpMaxVA, which becomes the upper-level optimization model, i.e.,

$$\text{Max } F(\mathbf{X},\mathbf{Z}) = \sigma(S,X,Y) = \sum_{i=1}^{N} z_i = \sum_{i=1}^{N}\sum_{j=1}^{J} z_{ij}, \tag{8.6a}$$

$$\text{s.t. } \Pr\left(P_j \middle| i,t\right) \geq \vartheta_i, \forall i \in [1,N], j \in [1,J], \tag{8.6b}$$

$$|S| = n, \tag{8.6c}$$

$$S \subseteq V. \tag{8.6d}$$

Here the design variables are the seed customers (i.e., S, corresponding to $\mathbf{X}$ in Eq. (8.3a)) and attribute levels (i.e., X and Y corresponding to $\mathbf{Z}$ in Eq. (8.3a)).

8.3.2 Lower-Level Optimization Model

For product portfolio planning, it is necessary to find an optimal combination of product attributes and attribute levels for the best product line performance in terms of utility-to-cost ratio. In this aspect, the lower level model considers both customer satisfaction and engineering cost so that the expected overall performance of the entire product line can be maximized. Among many, a shared surplus measure is selected. It leverages both customer satisfaction and engineering cost formulated as below (Jiao and Zhang, 2005b):

$$f(\mathbf{Z}) = \phi(X,Y) = \sum_{i=1}^{N}\sum_{j=1}^{J} \frac{U_{ij}}{C_j} z_i y_j, \tag{8.7}$$

U_{ij} is the utility of the *i*-th market segment (or *i*-th customer) for the *j*-th product. It is normally the case that the market can be grouped into several segments, in which all the customers are homogeneous and their utilities are similar. U_{ij} has a copulas formulation of part-worth utilities derived from cumulative prospect theory (CPT) of the attribute levels of product P_j in Chapter 5 (Zhou et al., 2014b), i.e.,

$$U_{ij} = C\left(u_{111} \dots, u_{ikl}, \dots, u_{IKL_K}\right) = c_1\varphi^{-1}\left[\prod_{k=1}^{K}\varphi\left(\xi_k + (1-\xi_k)\sum_{l=1}^{L_k} x_{jkl}\, u_{ikl}\right)\right] + c_2, \tag{8.8}$$

where C is a copulas function, u_{ikl} denotes the utility of the *l*-th attribute level of *k*-th attribute perceived by the *i*-th customer in the social network. u_{ikl} is often obtained by conjoint analysis. Here, it is derived from a customer preference model based on CPT (Zhou et al., 2014b). $c_1 = 1\Big/\left(1-\varphi^{-1}\left(\prod_{k=1}^{K}\varphi(\xi_k)\right)\right)$, $c_2 = 1 - c_1$, ξ_k is a constant, $0 \leq \xi_k < 1$, and φ is a generating function, which takes the form $\varphi(\xi_k) = \frac{1-\exp(\varsigma\xi_k)}{1-\exp(-\varsigma)}$, where $\varsigma \in R\backslash\{0\}$ (Zhou and Jiao, 2013b). C_j is a cost function of product P_j. We model the cost consequences of providing variety based on variation of process capabilities (Jiao and Tseng, 2004). A process capability index is an instrument for handling the sunk costs related to product lines and shared resources, based on which C_j is formulated as follows:

$$C_j = \varrho \exp\left(\frac{1}{PCI_j}\right) = \varrho \exp(\frac{3\sigma_j^T}{\mu_j^T - LSL^T}), \tag{8.9}$$

where ϱ is a constant indicating the average dollar cost per variation of process capabilities; LSL^T, μ_j^T, and σ_j^T are the lower specification limit, the mean, and the standard deviation of the estimated cycle time for product P_j. Finally, by maximizing $f(\mathbf{Z})$, I can obtain the set of product attributes and their levels, which can be represented as

$$\mathbf{Z}^* = argmax[f(\mathbf{Z})] = \{\mathbf{Z} | \forall \mathbf{Z}' : f(\mathbf{Z}') \leq f(\mathbf{Z})\}. \tag{8.10}$$

8.3.3 Bi-level Game Theoretic Optimization Model

By compiling the upper-level optimization model and the lower-level optimization model, I can obtain the bi-level optimization model as follows:

$$\text{Max } F(\mathbf{X}, \mathbf{Z}) = \sigma(S, X, Y) = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij}, \tag{8.11a}$$

$$\text{s.t. } z_{ij} = \frac{\text{sgn}[\Pr(P_j | i, t) - \vartheta_i] + 1}{2}, \forall i \in [1, N], j \in [1, J], \tag{8.11b}$$

$$\sum_{j=1}^{J} y_j \leq J^+, \tag{8.11c}$$

$$y_j, \in \{0,1\}, j \in [1, J], \tag{8.11d}$$

$$\sum_{l=1}^{L_k} x_{jkl} = 1, k \in [1, K], j \in [1, J], \tag{8.11e}$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L_k} |x_{jkl} - x_{j'kl}| > 0, j \neq j', \tag{8.11f}$$

$$x_{jkl} \in \{0,1\}, \forall i \in [1, N], j \in [1, J], k \in [1, K], l \in [1, L_k], \tag{8.11g}$$

$$|S| = n, \tag{8.11h}$$

$$S \subseteq V, \tag{8.11i}$$

$$\mathbf{Z}^* = argmax\left[f(\mathbf{Z}) = \phi(X, Y) = \sum_{i=1}^{N} \sum_{j=1}^{J} \frac{U_{ij}}{C_j} z_i y_j\right], \tag{8.11j}$$

$$\text{s.t. } U_{ij} = c_1 \varphi^{-1}\left[\prod_{k=1}^{K} \varphi\left(\xi_k + (1 - \xi_k) \sum_{l=1}^{L_k} x_{jkl} u_{ikl}\right)\right] + c_2 + \varepsilon_{ij}, \tag{8.11k}$$

$$C_j = \varrho \exp\left(\frac{1}{PCI_j}\right) = \varrho \exp\left(\frac{3\sigma_j^T}{\mu_j^T - LSL^T}\right), \tag{8.11l}$$

$$z_i = \sum_{j=1}^{J} z_{ij}, j \in [1, J], \tag{8.11m}$$

$$g_{jk}(X) \geq 0, j \in [1, J], k \in [1, K]. \tag{8.11n}$$

In this formulation, Eqs. (8.11a-8.11i) form the upper-level optimization model while Eqs. (8.11j-8.11n) form the lower-level optimization model. Eq. (8.11b) denotes the adoption condition, that is, if the adoption probability $\Pr(P_j|i,t)$ is larger than i-th customer's adoption threshold ϑ_i, $z_{ij} = 1$, otherwise $z_{ij} = 0$. Note in (8.11k), I add a random error parameter ε_{ij} for U_{ij}. Sgn is a sign function. Eq. (8.11n) describes the technical requirements and constraints of product j for the k attribute domains.

The leader controls the follower through product attributes X and product variants Y, indicating a priority for determining the product configuration. The follower, on the other hand, attempts to optimize a shared surplus $\phi(X,Y)$ between the customers and the producer subject to the decisions made by the leader. The follower returns an optimal configuration X^* and selected product variants Y^* to the leader, which influences U_{ij} in the upper level, and in turn z_{ij} in the upper-level's optimization model. Therefore, the bi-level optimization captures such interactions and coupling between the engineering domain and the marketing domain. The optimal solution (S^*, X^*, Y^*) thus indicates the 'best compromise' between these two domains, where S^* is the final set of customer seeds and X^* is the set of attribute levels that configure the product variants Y^*in the product line.

8.4 Model Solution

8.4.1 A Coordinate-Wise Optimization Strategy

The formulation in Eq. (8.11) is a mixed integer nonlinear bi-level programming problem which incorporates an InfMax problem. It actually involves two different kinds of variables, i.e., seed customers (S) and product attributes and variants (X, Y). In order to deal with this, I adopt a coordinate-wise optimization strategy based on (Barbieri and Bonchi, 2014) by decomposing it into two sub-problems and alternate the procedure of greedy seed selection and attribute update in the iterations of the search process. The greedy seed selection is solved with CELF++ (Goyal et al., 2011) due to its high efficiency, while the attribute update problem is solved by a HTGA (Li et al., 2014; Tsai et al., 2004). For the i-th iteration, improvement of the objective function $\sigma(S, X, Y)$ can be obtained, by forcing one of the following inequalities to hold:

$$\sigma(S^i, X^i, Y^i) \leq \sigma(S^{i+1}, X^i, Y^i), \tag{8.12a}$$

$$\sigma(S^i, X^i, Y^i) \leq \sigma(S^i, X^{i+1}, Y^{i+1}). \tag{8.12b}$$

When the product attributes and their levels are fixed, the greedy seed selection algorithm is applied to increase the size of the seed set S by making Eq. (8.12a) hold. When the seed set is fixed, the HTGA method is utilized to update the attribute levels and their corresponding product variants. It is observed that the seed set S does not relate to lower-level optimization problem directly. However, when the seed set S is fixed, both X and Y interact between the upper-level model and the lower-level model. Furthermore, σ is only a function of X and Y. By making use of the single-level parametric optimization problem introduced in Eq. (8.3), I can transform the bi-level optimization problem in Eq. (8.11) into a single-level form similar to Eq. (8.3), which is shown as follows:

$$\text{Max } F(X, \Psi(X)) = \sigma(X, \Psi(X)), \tag{8.13a}$$

$$\text{s.t.} \quad G'\big(X, \Psi(X)\big) \geq 0, X \in \{0,1\}^s, \tag{8.13b}$$

where Y is a unique response function of X, i.e., $Y = \Psi(X)$. Then Eq. (8.13) is solved with the proposed HTGA method. Therefore, the overall solution procedure is illustrated in Figure 8.2. Due to the fact that the InfMax problem has been well addressed in previous studies (Chen et al., 2009; Goyal et al., 2011), I mainly introduce the HTGA method that solves the bi-level optimization problem.

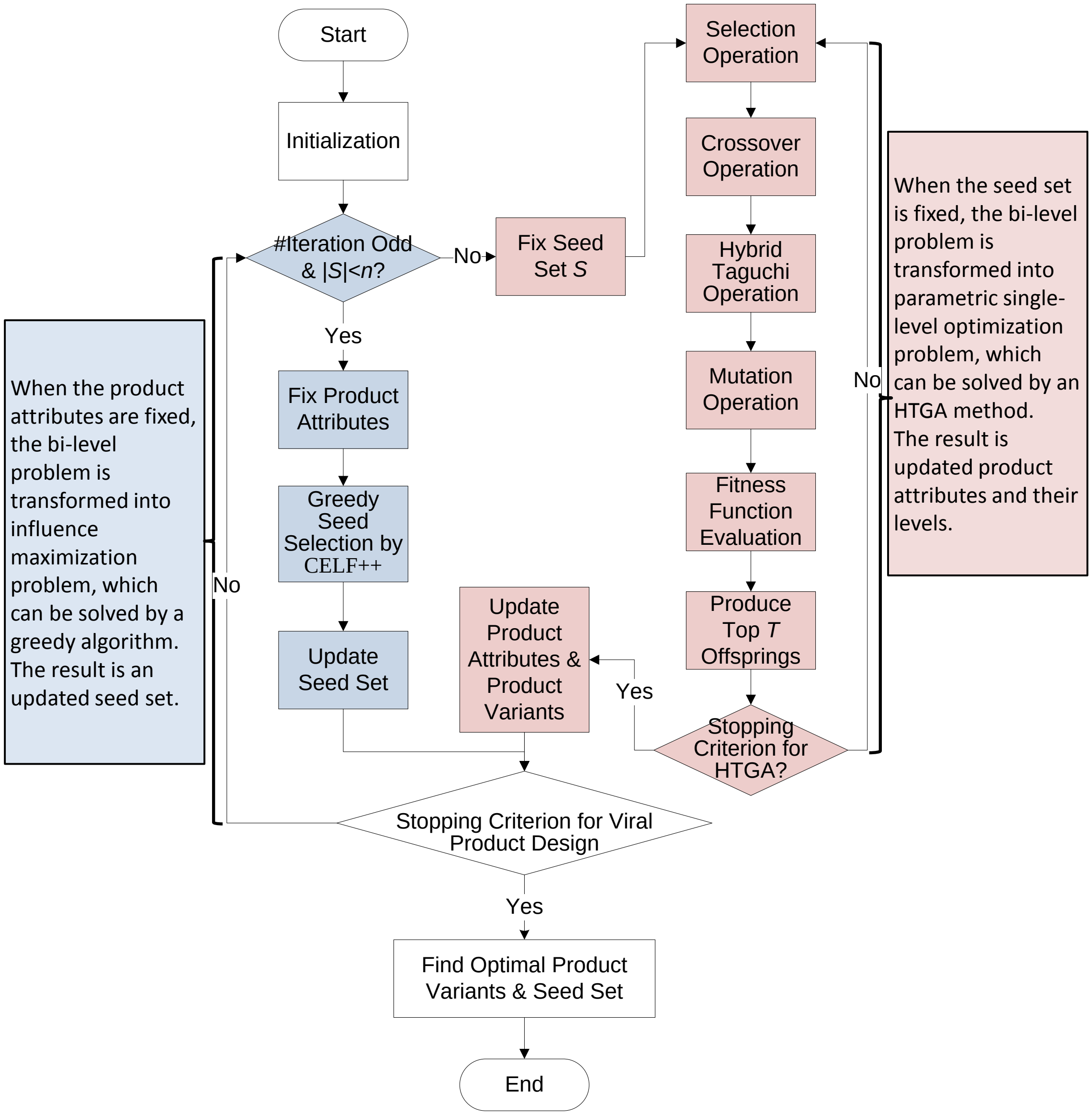


Figure 8.2 Solution procedure for the bi-level optimization model

8.4.2 Hybrid Taguchi-Genetic Algorithm

The HTGA method is based on orthogonal GAs in which some major steps (e.g., crossover) of a GA algorithm can be considered as "experiments" and the Taguchi's method or orthogonal design can be used to identify the global optimum more robust and statistically sound (Zhang and Leung, 1999). The Taguchi's method (Taguchi, 1995) can produce orthogonal arrays, i.e., fractional factorial design, which can scan the feasible space more evenly to locate good points for further exploration in the following iterations (Leung and Wang, 2001). Tsai et al. (2004) further introduce the notion of signal-to-noise ratio (SNR) to select better offspring, which further increases the quality and robustness of a GA, termed as the HTGA. In this research, I follow their idea and extend the original 2-level orthogonal arrays to n-level ($n > 2$) ones in the context of a bi-level optimization problem.

(1) Taguchi method: Both orthogonal arrays and SNR are used in the Taguchi method to generate high quality chromosomes in HTGA. Each product attribute is considered as a factor in the experiment design process, each attribute level is corresponding to a factor level. Usually multiple levels are associated with a factor. For example, the connectivity of a tablet can assume three levels, i.e., 4G LTE, WiFi, and both. Therefore, I extend the original two-level orthogonal arrays in (Tsai et al., 2004). An orthogonal array is a fractional factorial matrix with balanced comparison of factor levels and their interactions (Taguchi et al., 2000). For K factors with q levels per factor (when the numbers of factor levels are different, dummy variables can be used), the general symbol for a standard orthogonal array can be represented as

$$L_m(q^{m-1}) = [a_{i,j}]_{m\times(m-1)}, \tag{8.14}$$

where L denotes a Latin square, $m = q^k$is the number of experiment runs (i.e., number of rows in the orthogonal array), $k > 1$ is a positive integer, $m-1$ is the number of columns in the orthogonal array. Thus standard orthogonal arrays can be represented as $L_4(2^3)$ or $L_{27}(3^{26})$, for instance. They can be generated by algorithms (e.g., Leung and Wang, 2001) or software (e.g., SPSS). Table 8.1 shows part of $L_m(q^{m-1}) = L_{27}(3^{26})$ generated by SPSS, which is used in the case study. It has $K = 12$ factors (i.e., attributes) and each has two or three levels as shown in Table 8.2. Since $K \leq m-1$, only the first $K = 12$ columns are used, while the other $m-1-K = 14$ columns are ignored.

The notion of SNR is introduced to the Taguchi method in order to improve the robustness of the design. For example, turning the ignition key (signal factor) should always start the car, regardless of different noise factors, such as temperature, gasoline type, and engine wear. Three types of SNR tests exist based on the direction of the optimal path of the response variable (Taguchi et al., 2000) , i.e.,

Table 8.1 Orthogonal array for 12 product attributes of Kindle Fire HD tablets

#Exp.	Connectivity	Audio	Battery	Camera	EBook	Storage	Dimension	Price	Online Service	Video	Customer Service	Touch Screen
1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	1	1	3	2	1	2	3	2	1	1	1
3	2	1	3	1	1	1	2	2	1	1	2	1
4	1	2	3	2	3	1	1	3	2	1	1	1
5	3	1	2	2	2	3	1	1	1	1	1	1
6	2	1	1	2	1	2	1	2	1	1	1	2
7	1	2	3	3	1	3	2	1	1	1	1	2
8	3	1	2	1	1	1	1	3	2	2	1	2
9	3	2	1	1	2	2	2	1	2	1	2	2
10	1	2	3	1	2	2	1	2	1	2	1	1
11	3	1	3	2	1	2	2	3	1	2	1	1
12	2	1	3	3	3	2	1	1	2	2	2	1
13	2	2	2	1	2	2	1	3	1	1	1	1
14	1	1	1	3	3	2	1	3	1	1	1	2
15	3	1	3	1	3	3	1	2	2	1	1	2
16	3	2	1	2	3	1	1	2	1	1	2	1
17	1	1	1	2	2	3	2	2	2	2	1	1
18	1	1	2	3	2	1	1	2	1	2	2	2
19	2	2	2	3	1	3	1	2	2	1	1	1
20	2	1	1	1	3	3	1	1	1	2	1	1
21	1	1	2	2	1	2	1	1	2	1	2	1
22	2	2	2	2	3	1	2	1	1	2	1	2
23	3	1	2	3	3	2	2	2	1	1	1	1
24	3	1	3	3	2	1	1	1	1	1	1	1
25	2	1	3	2	2	3	1	3	1	1	2	2
26	1	1	2	1	3	3	2	3	1	1	2	1
27	3	2	1	3	1	3	1	3	1	2	2	1

Table 8.2 Product attributes and attribute levels identified for viral product design

Product Attributes	Connectivity	Audio	Battery	Camera	EBook	Storage	Dimension	Price	Online Service	Video Resolution	Customer Service	Touch Screen
Attribute levels	Wifi	Audible books	10hr	Front	<15000	16GB	7inch (395g)	199	Amazon Prime 1Yr	216ppi	Mayday button	Glare non-sensitive
	4G	Immersion reading	12hr	Back	<10000	32GB	8.9inch (545g)	249	Amazon Prime 2Yr	254ppi	No Mayday	Glare sensitive
	Wifi&4G	-	14hr	Front & Back	<5000	64GB	-	299	-	-	-	-

1) Nominal-is-best: for characteristics that need to be drawn as close as possible to a nominal value, e.g., outer diameter of a piston;

2) Smaller-the-better: for characteristics that need to have their response minimized, e.g., faulty rate in a production process;

3) Larger-the-better: for characteristics that need to have their response maximized, e.g., product adoption in a social network.

In this research, the larger-the-better case is adopted. Given a sample of n_o observations, $\{o_i | i = 1, \dots, n_o\}$, its SNR is calculated as follows:

$$\text{SNR} = -10\log\left(\frac{1}{n_o}\sum_{i=1}^{n_o}\frac{1}{o_i^2}\right), \tag{8.15}$$

(2) HTGA process: The HTGA process is shown in Figure 8.2. Before I introduce the HTGA process, the chromosome coding and the definition of the fitness function are described.

1) Chromosome coding: In Section 8.3.1, the J product configurations are represented as $X = \{X_1, \dots, X_J\}$, where the j-th product $X_j = \left(x_{j11}, \dots, x_{j1L_1}, \dots, x_{jK1}, \dots, x_{jKL_K}\right)$, and $x_{jkl} = 1$ if the *l*-th attribute level of the *k*-th attribute is selected for P_j, otherwise it is 0. And $Y = \left(y_j | j = 1, \dots, J\right)$ is a vector indicating a particular choice of P_j, and $y_j = 1$ if P_j is chosen, otherwise it is 0. However, I code a chromosome with only one product configuration at a time, i.e., X_j. Furthermore, based on Table 8.2, I transform the binary string of X_j into $X_j = \{x_{j1}, x_{j2}, \dots, x_{jK}\}$, where $x_{jk} \in \{1, 2, 3\}$, representing the specific level chosen by P_j. First, this will dramatically reduce the length of the binary string, which increases the efficiency of the HTGA process. Second, I will choose the best J^+ product configurations based on the seed set S and attribute level set A^*, and these product configurations actually determine the value of Y. Therefore, a chromosome c_j of an upper-level decision variable X_j can be represented as a string, i.e., $c_j = X_j = \{x_{j1}, x_{j2}, \dots, x_{jK}\}$, where $x_{jk} \in \{1, 2, 3\}$ (see Table 8.2).

2) Fitness function: I propose to apply an exact penalty function to handle constraints involved in Eq. (8.13) (Li et al., 2014), which can be represented as follows:

$$fit(c_j) = \begin{cases} \sigma\left(c_j, \Psi(c_j)\right) + p\left(c_j, \Psi(c_j)\right), \Psi(c_j) \text{ exists} \\ -\infty, \text{otherwise} \end{cases}, \tag{8.16}$$

where $p\left(c_j, \Psi(c_j)\right) = C_p min\left(G'\left(c_j, \Psi(c_j)\right), 0\right)$, where C_p is the penalty constant and takes a large positive value. It shows that when $G'\left(c_j, \Psi(c_j)\right) < 0$, $p\left(c_j, \Psi(c_j)\right)$ will have a small negative value, which substantially reduces the value of fitness function.

3) Initialization: By making use of the pseudorandom uniformly distributed function *rand* in Matlab, I first generate a uniformly distributed string with length K, i.e., $r = \{r_1, \dots, r_K\}$. If $r_k \in [0,1/3]$, I set $x_{jk} = 1$, and $r_k \in (1/3,2/3]$, I set $x_{jk} = 2$, otherwise, $x_{jk} = 3$ such that a random chromosome can be produced. This process is repeated for M_p times, which forms the initial population with M_p randomly generated chromosomes.

4) Selection: In the first iteration, the algorithm keeps all the M_p chromosomes. However, in the following iterations, the algorithm has a selection process, in which the fitness function is evaluated and ranked in a descending order. Then the top M_p chromosomes will be kept. This is due to the fact that in the following process of HTGA (e.g., crossover and mutation), more chromosomes will be generated.

5) Crossover: It is a genetic operator that can vary the programming of a chromosome from one generation to the next. There are multiple crossover techniques exist, such as one-point crossover, two-point crossover, cut and slice, and uniform crossover, and so on (Holland, 1992). In this research, the simple but effective one-point crossover technique is used with a crossover probability p_c. For two randomly selected parents $c_1 = \{x_{11}, \dots, x_{1i}, x_{1,i+1}, \dots, x_{1K}\}$, $c_2 = \{x_{21}, \dots, x_{2i}, x_{2,i+1}, \dots, x_{2K}\}$, it randomly selects the i-th cut-point, and exchanges the right parts of two parents to form two new generations, $c_1' = \{x_{11}, \dots, x_{1i}, x_{2,i+1}, \dots, x_{2K}\}$, $c_2' = \left\{x_{21}, \dots, x_{2i}, x_{1,i+1}, \dots, x_{1K}\right\}$. This process will repeat for M_p times.

6) Hybrid Taguchi operation: Both orthogonal arrays and SNR are used in this operation. In order to illustrate the process of the hybrid Taguchi operation, I use the example in Table 8.1. In this example, there are 12 factors involved in the product Kindle Fire HD tablet. An orthogonal array $L_{27}(3^{26})$ is produced with only the first 12 columns, in which each factor has 2 or 3 levels as shown in Table 8.2. Each row forms a configuration of the product, which is corresponding to X_j. For example, for the first factor connectivity, it has three levels ($L_1 = 3$), WiFi, 4G, and both, indicated as '1', '2', and '3' in Table 8.1. If WiFi is selected in the first row, then $x_{11} = 1$. For the purpose of illustration, I want to identify an optimal product configuration that can maximize a simplified version of the shared surplus in Eq. (8.17),

$$\phi = \sum_{k=1}^{12} \frac{u_k}{\pi_k}, \tag{8.17}$$

where the part-worth utility u_k is derived via a customer preference model based on CPT (Zhou et al., 2014b) and its cost π_k is derived with Eq. (8.9). Their values are shown in Table 8.3. We assume that three chromosomes are selected to determine which level is the best one for each factor, i.e., $c_1 = \{2,1,3,1,3,1,2,1,1,2,3,1\}$, $c_2 = \{3,2,1,3,2,2,1,2,2,1,1,2\}$, and $c_3 = \{1,3,2,2,1,3,3,3,3,3,2,3\}$. Then the values in c_1, c_2, and c_3 are considered as the new level 1, level 2, and level 3, respectively. For example, for battery, level 1, level 2, and level 3 are now 3, 1, and 2, respectively. By filling in Table 8.1 with these new levels, I can calculate their function values using ϕ in Eq. (8.17) from the first experiment to the 27th experiment. Define the effects of factors as

$$E_{fl} = \text{sum of } \mathrm{SNR}_i \text{ for factor } f \text{ at level } l, \tag{8.18}$$

where SNR_i is the SNR for the i-th experiment. In Table 8.1, I first calculate $\mathrm{SNR}_i, i = 1, \ldots, 27$ for the 27 experiments according to Eq. (8.15). Then Eq. (8.18) is used to compute $E_{fl}, f = 1, \ldots, 12, l = 1, 2, 3$. By comparing the values among E_{f1}, E_{f2}, and E_{f3}, the largest value corresponds to the optimal level based on the larger-the-better criterion (see Table 8.3). Note in the optimal level row, $c_2(1)$ denotes the first gene of the chromosome c_2, which is 3, i.e., the first gene of the produced chromosome. Thus, the final produced chromosome is $c_p = \{3,1,1,1,2,3,2,2,1,2,3,1\}$. Using Eq. (8.17), I can obtain their shared surplus of the four product configurations corresponding to the four chromosomes, $\phi(c_1) = 8.32$, $\phi(c_2) = 10.75$, $\phi(c_3) = 5.38$, and $\phi(c_p) = 11.26$. It validates that the produced chromosome is the optimal one among the four. Thus, this hybrid Taguchi operation can produce better offspring with three parents without computing all the possible product configurations (i.e., $3^6 \times 2^6 = 46656$), but only 27 orthogonal configurations. This process can greatly improve the convergence rate of the proposed GA method, especially important for the complex bi-level optimization problem in this research.

7) Mutation: It is a genetic operator that is used to maintain and introduce genetic diversity from one generation to the next. Multiple mutation techniques exist, such as flip bit, boundary, uniform or non-uniform, and Gaussian, and so on (Holland, 1992). In this research, the algorithm first chooses a chromosome randomly from those generated in the previous steps with probability p_m. Then two genes randomly selected from the chromosome will be altered from one level to another level (e.g., from level 1 to level 2 or level 3 with an equal probability). This process will be repeated for M_p times.

8) Fitness function evaluation: All the chromosomes, including those produced by crossover, the hybrid Taguchi operation, and mutation, will be evaluated in terms of their fitness using the fitness function in Eq. (8.16). Then they will rank in a descending order and the top M_p chromosomes will be kept as the new population for the next iteration until the stopping criterion is met.

9) Stopping criteria: It is controlled by a maximum iteration number I_{max}. However, when the iteration number is smaller than I_{max}, I also detect whether the algorithm converge after a minimum iteration number I_{min} by comparing the optimal fitness function values at the current iteration and at the previous one. If I detect that the optimal fitness function value does not change for over a minimum number of times C_{min}, I consider that the algorithm has converged. Then I will stop the algorithm also.

Table 8.3 Optimal chromosome generation process based on the hybrid Taguchi operation

Factor	Connectivity	Audio	Battery	Camera	EBook	Storage	Dimension	Price	Online Service	Video Resolution	Customer Service	Touch Screen
Utility /Cost	0.14/ 0.19	0.29/ 0.28	0.15/ 0.11	0.16/ 0.13	0.58/ 0.65	0.05/ 0.19	0.5/0.49	-0.11/ 0.21	0.34/0.4	0.36/0.44	0.98/ 0.99	0.98/0.84
	0.23/ 0.26	0.71/ 0.72	0.29/ 0.25	0.17/ 0.15	0.38/ 0.24	0.42/ 0.35	0.5/0.51	-0.16/ 0.36	0.66/0.6	0.64/0.56	0.02/ 0.01	0.02/0.14
	0.63/ 0.55	0/1000	0.56/ 0.64	0.67/ 0.72	0.04/ 0.11	0.53 /0.46	0/1000	-0.73/ 0.43	0	0/1000	0	0
$c_1: E_{f1}$	397.8	**797.2**	393.9	**400.8**	386.2	386.2	**798.2**	405.6	**794.5**	**802.2**	**783.5**	**810.6**
$c_2: E_{f2}$	**403.3**	398.1	**402.3**	395.2	**411.7**	404.4	397.2	**408.0**	400.8	393.1	411.8	384.7
$c_3: E_{f3}$	394.3	0.0	399.1	399.4	397.5	**404.8**	0.0	381.7	0.0	0.0	0.0	0.0
Optimal level	$c_2(1)$	$c_1(2)$	$c_2(3)$	$c_1(4)$	$c_2(5)$	$c_3(6)$	$c_1(7)$	$c_2(8)$	$c_1(9)$	$c_1(10)$	$c_1(11)$	$c_1(12)$
Produced chromosome	3	1	1	1	2	3	2	2	1	2	3	1

Note that for those attributes which do not have a third level, their utilities are 0 and costs are indicated with big numbers (i.e., 1000). This trick prevents the algorithm from choosing them in the hybrid Taguchi operation. In the row of optimal level $c_2(1)$ is the first gene of the chromosome c_2. Those in bold are the best among the three levels, and thus are chosen for the produced chromosome in the table.

8.5. *Case Study*

8.5.1 Data Collection

A case study of Kindle Fire HD tablets (released in 2012) from Amazon.com is used to illustrate the proposed method. Usually after a review is made by a customer, others will comment on the review after reading it. Thus, links from the reviewer to the commenters are constructed and potential influence from the reviewer to the commenters can take place. Based on this idea, reviews of Kindle Fire HD tablets 7 inch tablets are collected from September 2012 to September 2013 (50 weeks), and the social network is constructed based on the reviewer-commenter links. The social network has 5220 nodes and 10476 edges. As mentioned in Chapter 7, the degree distribution on a log-log scale resembles a power law distribution, and it has a giant connected component along with a large number (i.e., 244) of smaller connected components. Hence, the social network generated is consistent with past studies on large-scale social networks (Leskovec and Horvitz, 2008).

A total number of 12 product attributes are identified by sentiment analysis in Chapter 5 as shown in Table 8.2 and each has two or three levels, which are related to both popularity and latent customer needs that often delight customers unexpectedly. We assume that latent customer needs satisfied by certain attributes possibly contribute to product adoption in the social network. The case study is to illustrate the proposed bi-level optimization method as a way to evaluate viral product design by identifying which product configuration is the optimal one in terms of product adoption and shared surplus maximization under the influence of social network effects.

8.5.2 Data Analysis

Based on the collected data, the input of the bi-level optimization includes

(1) The social network constructed from the reviewer-commenter links about Kindle Fire HD tablets on Amazon.com $G = (V, E)$, where $V = \{v_1, v_2, \ldots, v_N\}, N = 5220$ is a set of nodes, indicating a set of customers and a link from v_j to v_i, i.e., $(v_j, v_i) \in E$ means that v_j can potentially influence v_i. The total number of links is 10467;

(2) A set of product attributes $A = \{a_1, \dots, a_K\}, K = 12$ and a set of their attribute levels in Table 8.2, i.e., $A^* = \{a_{kl}^*\}, k = 1, \dots, 12, l = 1, \dots, L_k, L_k = 2$ or 3;

(3) The budget indicated as a set of seed customers $S,\ |S| = n < 5220$. I choose different values of n for experimentation;

(4) The diffusion model in Eq. (8.6) and its parameters, including peer influence on customer v_i, i.e., I_i^t , the utility of the i-th market segment (or i-th customer) for the j-th product, U_{ij}, and customer v_i's hurdle utility for the j-th product, h_{ij}, indicating equilibrium between adopting and not adopting the product. However, I need to estimate these parameters before I can run the optimization model;

Based on the previous chapter, I can compute I_i^t based on three operational factors of the peer influence, including interaction strength, structural equivalence, and entity similarity (Zhou et al., 2014c).

However, it is not practical to compute U_{ij} for every customer in the social network. Nevertheless, it is possible to segment the market based on the m_c customer characteristics from their profiles on Amazon.com in Chapter 7. Using a data mining method named rough set (Pawlak, 1991), I identify five market segments based on the adoption categories, including innovators, early adopters, early majority, late majority, and laggards (Rogers, 2003). Then all the customers in the same i-th market segment are considered as homogeneous and their utility of the j-th product configuration is $U_{ij} + \varepsilon_{ij}, i = 1, \dots, 5$. In such a way, I can compute U_{ij} with the method in Chapter 6 (Zhou et al., 2014b). ε_{ij} is a random error term for each segment-product pair and it follows a normal distribution $\varepsilon_{ij} \sim N(0,10)$. The standard deviation is estimated based on the difference within market segments. I assume that customers within a market segment are more homogeneous than those between different market segments. Therefore, the variance of between segments of ε_{ij} is estimated around 4 times that of within segments.

In order to estimate the hurdle utility h_{ij}, a maximum likelihood method is used (Barbieri and Bonchi, 2014). Based on the reviews posted on Amazon.com, I can have a log of past product adoption for all the customers, denoted as $D^l = (V, P, T)$, in which each tuple $\langle i, P_j, t \rangle$ indicates customer v_i adopted the product P_j at time t. Let Θ denote the set of all the parameters, i.e., hurdle utilities that need to estimate. The likelihood of the data maximization problem given the model parameters H can be expressed as $L(D^l;\Theta) = \sum_{P_j \in P} log\left(L(D_j^l;\Theta)\right)$, subject to the universe of products P. Let D_j^l represent the propagation of P_j (i.e., corresponding to tuples as long as the product is P_j). Then $L(D_j^l;\Theta) = \prod_{v_i \in V} \Pr(P_j|i,t)^{\delta(i,P_j)}\left(1-\Pr(P_j|i,t)\right)^{1-\delta(i,P_j)}$, where t is the time when v_i adopted P_j, and $\delta(i,P_j) = 1$ if v_i adopted P_j, and 0 otherwise. $\Pr(P_j|i,t)$ is shown in Eq. (8.4). Then the estimation of h_{ij} can be solved by the following optimization problem, i.e.,

$$\text{Max}_H\ log\ L(D^l;\Theta), \tag{8.19a}$$

$$\text{s. t. } h_{ij} > 0, v_i \in V, P_j \in P. \tag{8.19b}$$

(5) Initialization parameters for the HTGA method, including population size $M_p = 1000$, crossover probability $p_c = 0.8$, mutation probability $p_m = 0.2$, an orthogonal array $L_{27}(3^{26})$ with only the first 12 columns as shown in Table 8.1, penalty parameter $C_p = 10000$, the maximum number of iteration $I_{max} = 1000$, the minimum number of iteration $I_{min} = 50$, and the minimum number of convergence $C_{min} = 10$.

(6) To calculate C_j in Eq. (8.6), I need to estimate ϱ and LSL^T. According to Jiao and Tseng (2004), these parameters can be estimated from empirical studies and I assume that $\varrho = 180$ and $LSL^T = 500$ seconds.

8.5.3 Results and Validation

The output of the bi-level optimization model includes a set of seed customers $S, |S| = n$ and a set of J^+product configurations that maximize product adoption at the upper level and maximize shared surplus at the lower level. By adopting the strategy in Eq. (8.12), iteration by iteration, I utilize CELF++ (Goyal et al., 2011) (code available at www.cs.ubc.ca/~goyal/code-release.php) to identify S due to its efficiency and flexibility in selecting diffusion models and adopt HTGA to identify the optimal attribute levels due to its fast convergence. In order to validate the proposed model, both viral product attributes and viral influence attributes are identified. The contributions of viral product attributes (in terms of the expected number of product adopters and the expected shared surplus (see Figure 8.5)) are demonstrated by comparing the results with those obtained by methods used in viral marketing. Moreover, by comparing the selected seeds with all the social entities in the social network in terms of multiple measures (see Table 8.5), the identified viral influence attributes are proved to be more important and influential than all of the social entities as a whole.

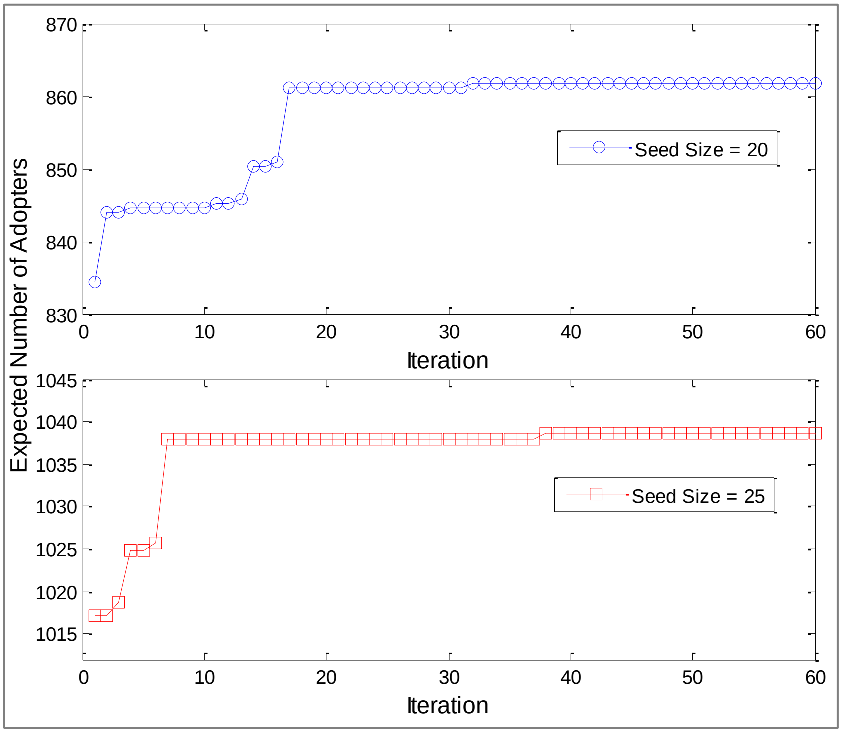


Figure 8.3 Convergence for the upper-level model

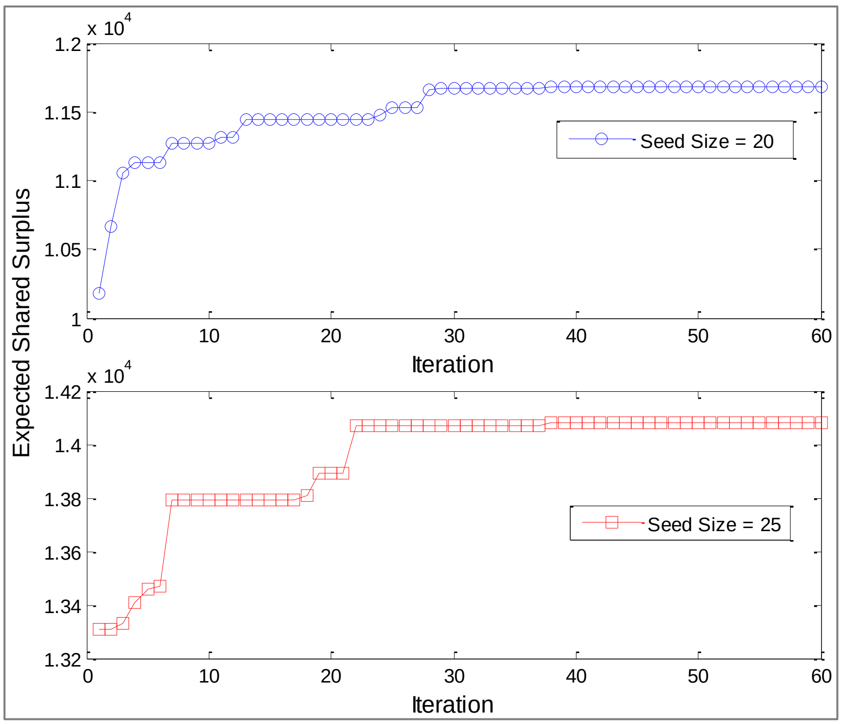


Figure 8.4 Convergence for the lower-level model

Table 8.4 Optimal product configurations identified with the proposed method

Seed size	Optimal chromosome ($J^+ = 3$) or Optimal product configuration	Expected number of adopters	Expected shared surplus
$\lvert S\rvert = 20$	c_{opt1} = {3,1,1,1,2,2,2,2,2,2,2,1}: Wifi&4G, Audible books, 10hr, Front, <10000, 32GB, 8.9inch, 249, Amazon Prime 2Yr, 254ppi, No Mayday, Glare non-sensitive	862	11680
	c_{opt2} = {3,1,1,2,2,2,1,1,2,2,2,1}: Wifi&4G, Audible books, 10hr, Back, <10000, 32GB, 7inch, 199, Amazon Prime 2Yr, 254ppi, No Mayday, Glare non-sensitive	861	11589
	c_{opt3} = {2,2,1,2,2,2,2,2,2,2,2,1}: 4G, Immersion reading, 10hr, Back, <10000, 32GB, 8.9inch, 249, Amazon Prime 2Yr, 254ppi, No Mayday, Glare non-sensitive	845	11112
$\lvert S\rvert = 25$	See c_{opt1} when $\lvert S\rvert = 20$	1039	14078
	See c_{opt2} when $\lvert S\rvert = 20$	1038	13971
	c_{opt3} = {2,1,1,1,2,2,2,2,2,2,2,1}: 4G, Audible books, 10hr, Front, <10000, 32GB, 8.9inch, 249, Amazon Prime 2Yr, 254ppi, No Mayday, Glare non-sensitive	1026	13636

I set $|S| = 20$ and 25 to test the model. Figure 8.3 and Figure 8.4 show the convergence of the algorithm in terms of both the lower-level model and the upper-level model. From Figure 8.3, I can tell the maximum numbers of adopters for $|S| = 20$ and 25 are 862 and 1039, respectively. Similarly, from Figure 8.4, the maximum numbers of shared surplus for $|S| = 20$ and 25 are 11680 and 14078, respectively. Note that shared surplus is proportional to the number of adopters. Both figures show the convergence in a relatively small numbers. On the one hand, the fast convergence shows the efficiency of the HTGA method, and each iteration shows the strategy in Eq. (8.12), which incorporates the CELF++ algorithm. On the other hand, it shows the equilibrium between the leader and the follower. The top three optimal chromosomes derived are shown in Table 8.4. First, the top 2 optimal product configurations are the same for two different seed sizes (i.e., $|S| = 20$ and 25) while the third one is different. To some extent, this shows that some product attributes consistently contribute to product adoption regardless of the influence of social network effects. These product attributes form the set of viral product attributes, including 10hr, <10000, 32GB, Amazon Prime, 2Yr, 254ppi, No Mayday, Glare non-sensitive. Within the same seed size, the differences (shown in bold in Table 8.4) in product configurations are manifested themselves in terms of connectivity, audio, camera, dimension, and price. For example, for the top 2 product configurations, the main differences are camera, tablet dimension, and price. This seems to reflect the major product variants provided in the market (i.e., Kindle Fire HD 7 inch and 8.9 inch tablets). However, one level seems to dominate another for other product attributes (e.g., video resolution, storage, online service, battery, and customer service). For example, the storage is always 32GB for the top 3 product configurations. This seems inconsistent with the current product variants in the market.

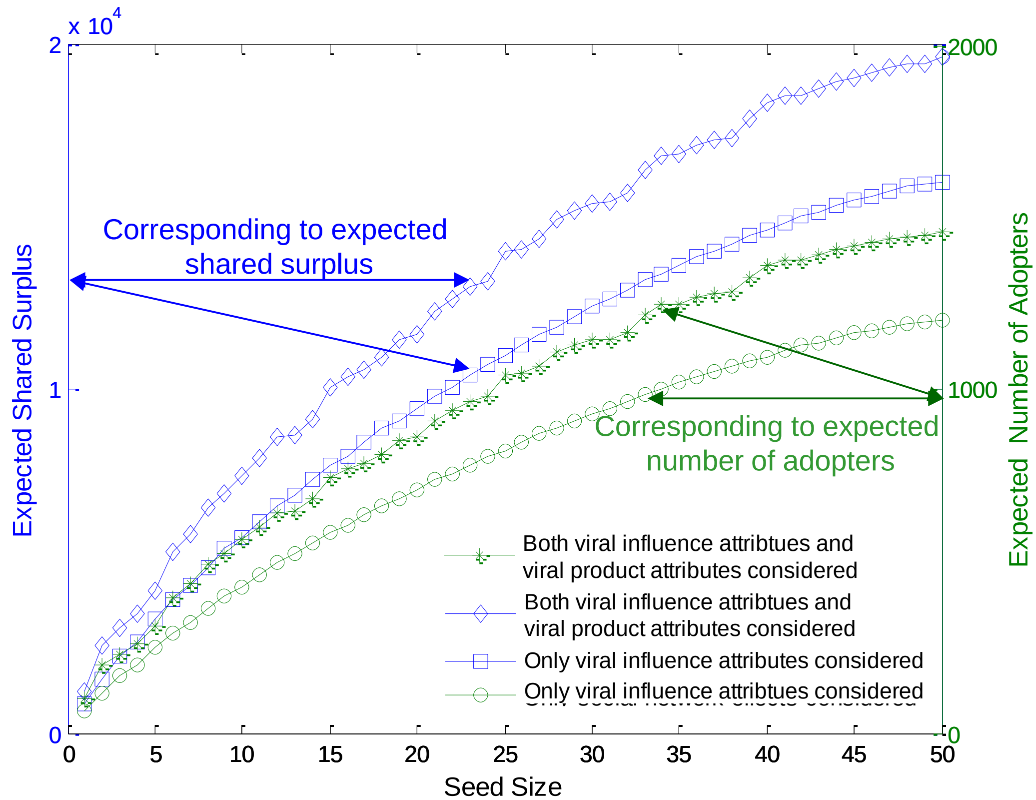


Figure 8.5 Influence of the seed (viral influence attributes) size and viral product attributes on product adoption and shared surplus

Furthermore, I also study the influence of the seed (i.e., viral influence attributes) size and viral product attributes on product adoption and shared surplus as shown in Figure 8.5. It has one x axis, i.e., seed size and two y axes, in which the left one is the expected shared surplus (in blue) and the right one is the expected number of adopter (in green). For one thing, when only viral influence attributes are considered, both the expected shared surplus and the expected number of adopters increase steadily when the seed size increases. It can be seen that these two curves have properties of both monotonicity and submodularity. When both viral influence attributes and viral product attributes are considered, the expected shares surplus and the expected number of adopters also increase. However, it seems that submodularity is not expected in these two curves. Some larger marginal is shown when the seed size is larger than when it is smaller. It may be due to the fact that part-worth utilities of some attribute levels may be positive for some customers, but may be negative for others as well. In this aspect, the strategy in Eq. (8.12) and the proposed HTGA method helps solve the bi-level optimization problem efficiently. For another, I can see the gap between two blue curves and that between two green curves seem to increase as the seed size increases. This shows that the contribution of viral product attributes to product adoption.

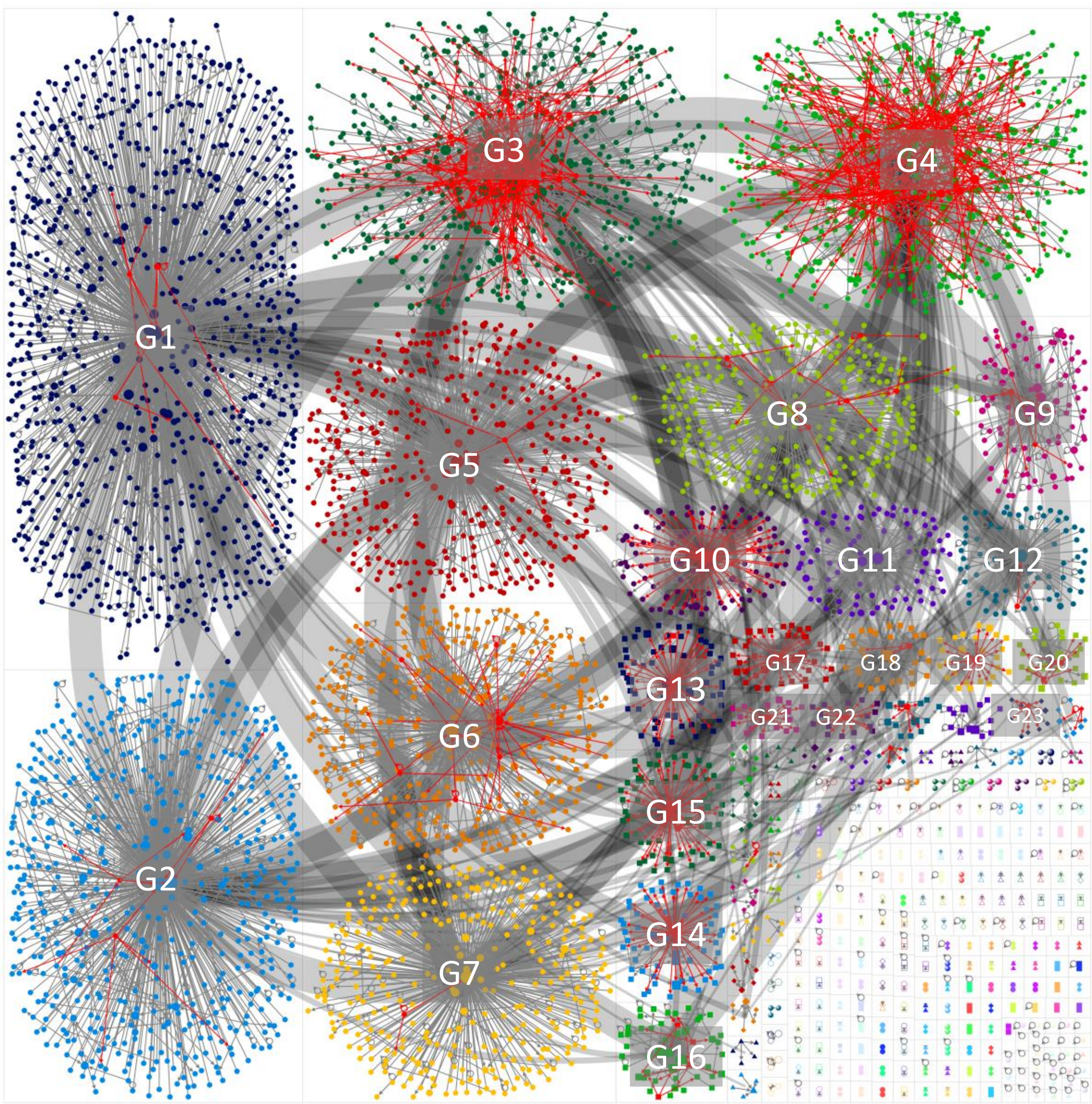


Figure 8.6 Adopters resulting from 50 selected seeds in the social network

Figure 8.6 shows a total number of 1459 adopters in the social network caused by the influence of 50 selected seeds (i.e., viral influence attributes) in Figure 8.5. The figure is generated by NodeXL (http://nodexl.codeplex.com) with the Fruchterman-Reingold layout (Fruchterman and Reingold, 1991) in the form of community structures obtained by the Clauset-Newman-Moore clustering algorithm (Clauset et al., 2004). Only larger groups are labeled with their numbers in the social network. The red nodes show the seeds and the red edges link to the activated adopters by these 50 seeds, mainly in G3, G4, G10, G13, G14, G15, and G17. Close examination shows that G1, G2, G5, G6, G7, and G8 are centered on nodes which have larger in degrees. G3 and G4 are mixed in which both larger in-degree and out-degree nodes exist. Most seeds have relatively larger out degrees in G10, G13, G14, and G15. Many adopters (i.e., commenters) actually have multiple rounds of interactions with one seed customer (i.e., reviewer), which tends to increase the influence from the reviewer to the commenter.

In order to further understand the seed customers selected, I analyze the statistics of 50 seeds selected in Figure 8.5 and compare them with all the social entities in the social network as a whole. The results are shown in Table 8.5. Seven measures are included. Ratings and review helpfulness are relevant to the review content, while other measures are derived from the social network with NodeXL. Ratings on Amazon.com are based on a 5-point Likert scale, where 1 indicates the review is most negative and 5 indicates the review is most positive. From Table 8.5, I can tell that both the selected 50 seeds and all the social entities positively evaluate the product. However, the minimum rating and the mean rating of the selected seeds are larger than those of all the social entities. This shows that the positive reviews tend to influence others in the social network to adopt the product. The measure of review helpfulness is defined as the ratio of n_h/n_a that n_h out of n_a Amazon customers find the review helpful. It shows that the selected seeds provide more helpful reviews than all the social entities in general.

Centrality measures, including degree, betweenness, closeness, eigenvector, and PageRank, tend to measure how important and influential a node is in a social network, although some disagree (Tsvetovat and Kouznetsov, 2011). Out-degree shows how many people commented on a social entity's review or how many people the review directly influences others in the social network. From Table 8.5, the selected seeds tend to have more comments statistically than all the social entities, especially for the mean and median values. Betweenness centrality indicates the number of times a node acts as a bridge along the shortest path between two other nodes in the social network (Newman, 2005). Hence, those with high betweenness tend to bridge different communities in the social networks. Closeness is the inverse of farness which is defined as the sum of one node's distances to all the other nodes. Thus the more central a node is, the more close it is to all other nodes, which can be interpreted as how long it will take to spread information from the node to others sequentially (Newman, 2005). Eigenvector centrality can be considered as a weighted sum of both direct connections and indirect connections of every length of one node with regard to other nodes in a social network (Bonacich, 2007). The weights are calculated based on nodes centrality and thus it takes the entire network pattern into account. PageRank is determined by the incoming links of a node and represents the likelihood that other social entities will encounter the node through the links provided (Tsvetovat and Kouznetsov, 2011). As for the centrality measures shown in Table 8.5, although the maximum values of the selected seeds are smaller than those of all the social entities, the minimum, the median and the mean values are larger than those of all the social entities. Therefore, from both the review content and importance and influence of nodes in the social network, the selected seeds tend to be of 'high quality' than all the social entities as a whole.

Table 8.5 Statistic comparisons of multiple measures between 50 selected seeds and all the social entities

Measures	50 Selected Seeds				All Social Entities			
	Max	Min	Median	Mean	Max	Min	Median	Mean
Ratings	5	3	5	4.5	5	1	5	4.1
Review helpfulness	100%	44%	62%	64%	100%	0%	54%	54%
Out-degree	746	8	10	25	746	0	1	1.7
Betweenness	662237	8661	70054	128424	1.3e7	0	0	12880
Closeness	7.6e-5	4.6e-5	6.1e-5	6.2e-5	1.0	0	0	5.6e-5
Eigenvector	1.0e-3	6.4e-5	1.7e-4	2.8e-4	0.2e-1	0	0	2.5e-4
PageRank	27.1	1.8	2.6	5.0	337.4	0.3	0.6	1.0

8.6 Discussions

(1) Viral product design: The concept of viral product design integrates product portfolio planning and viral marketing. From the product portfolio planning point of view, it aims to identify an optimal set of viral product attributes and attribute levels in order to maximize product adoption and its performance, such as shared surplus which considers both customer satisfaction and engineering cost. From the viral marketing point of view, it accommodates both viral influence attributes and viral product attributes so as to maximize product adoption. By integrating the SoC problem in product portfolio planning and InfMax in viral marketing, this chapter proposes the AdpMaxVA problem. In order to tackle both product portfolio planning (maximize product adoption and product line performance) and viral marketing (maximize product adoption), I propose a bi-level optimization model, in which the AdpMaxVA problem is modeled as the leader and the problem of maximizing product line performance is modeled as the follower. Such a formulation bridges the gaps between engineering design and marketing in the context of social networks, and the interplay and coupling between product design and viral marketing are captured within a coherent framework of joint optimization.

(2) Solution to bi-level optimization: In order to solve the bi-level optimization problem, I follow a coordinate-wise maximization strategy as shown in Eq. (8.12). It allows us to alternate local greedy choices for the update of the seed set and the attribute set (Barbieri and Bonchi, 2014). When the set of product attributes is fixed, it becomes an InfMax problem. By capitalizing on the monotonicity and submodularity of the adoption function σ, I adopt the CELF++ algorithm (Goyal et al., 2011). When the seed set is fixed, I propose a HTGA method. First, I transform the bi-level problem into a single-level parametric optimization problem. This transformation guides us to the solution with a GA, in which the fitness function can be obtained with a penalty function. Second, the proposed HTGA method makes use of the Taguchi method in experiment design to improve the robustness and efficiency in the chromosome generation process. The chromosome is coded as one particular product configuration in terms of attribute levels rather than a binary string. It not only avoids the constraints in Eq. (8.11e) but also improves the efficiency of the algorithm. As shown in Figure 8.3 and Figure 8.4, fast convergence can be obtained in a relatively small number of iterations.

(3) Implications: First, from the top 3 optimal product configurations obtained for two different seed sizes, it shows that product variants only manifest themselves in a small number of product attributes. This result follows the practice in product portfolio planning and mass customization. By only changing a small number of product attributes in terms of their attribute levels, the cost in manufacturing can be reduced substantially. This is consistent with the objective in the lower-level optimization problem, i.e., shared surplus considering both customer satisfaction and engineering costs.

Second, by always keeping some of the product attribute levels, such as storage being no smaller than 32GB and the touch screen being glare non-sensitive, the product can satisfy latent customer needs and maximize product adoption at the same time. As mentioned previously, the selection of these product attributes and attribute level is actually derived from Chapter 5 on latent customer needs elicitation based on the product reviews about Kindle Fire HD tablets from Amazon.com. The 16GB storage has received many negative reviews while one of the latent customer needs is that the touch screen is glare non-sensitive. Therefore, the optimal product configurations obtained in terms of maximum product adoption can be attributed to customer satisfaction and latent customer needs to a large extent.

Third, Figure 8.5 proves that both viral product attributes and viral influence attributes are conducive to product adoption in the context of social networks. By engineering viral product attributes in the design process and by making use of the social interactions between customers, especially in the online social networks without geographic constraints, product diffusion and adoption can be facilitated and catalyzed to the largest possible extent.

Fourth, InfMax algorithms identify the optimal seed set by considering both the diffusion mechanism and the social network structure. Particularly when the social network structure is given, it is important to leverage review content for adoption maximization. The proposed diffusion model in Eq. (8.6) includes both peer influence, customer preferences in terms of holistic product utilities, and individual hurdle utilities. The results in Figure 8.6 and Table 8.5 generally coincide with the proposed diffusion model and the InfMax algorithm. Therefore, it is possible that firms can enhance their viral marketing strategy by better identifying the seeds and promoting the product in social media. Moreover, many social network websites provide basic profiles about customers, such as social, demographic, and behavioral information. By making use of such information and then personalizing incentives to different seed customers according to their influence and importance in the social network, companies may better promote social commerce in the cyber environment.

8.7 Summary

The proposed bi-level game theoretic optimization model combines engineering design, social network analysis, and viral marketing in a coherent fashion. Such integration of viral influence attributes in viral marketing and viral product attributes in product design is a significant contribution to the amelioration of decision analysis in design for market systems, which also sheds light on understanding the social aspect of design at large.

CHAPTER 9

CONCLUSIONS AND FUTURE WORK

This concluding chapter summarizes the findings and the contributions of the thesis work. The limitations and possible improvements are also discussed, along with avenues for future research.

9.1 Conclusions

Social media have fundamentally transformed the way we communicate, collaborate, consume and create (Aral et al., 2013). These tools have provided firms with unprecedented capabilities to reach and capitalize on the network users and what they create on these social networks. While peer influence of social networks has been widely used in marketing related activities, limited efforts have been paid attention to design. In this aspect, I propose a new paradigm of design, i.e., viral product design for social network effects. It makes use of peer influence of social networks by identifying viral product attributes and viral influence attributes for maximizing product adoption and optimizing product line performance. In order to tackle such a complex design problem systematically, a technical framework is proposed, dealing with issues of latent customer needs elicitation, customer preference modeling and quantification, social network modeling, and viral product design evaluation.

First, latent customer needs elicitation aims to extract potential viral product attributes that can delight customers unexpectedly. Unlike traditional customer needs elicitation, I propose use case analogical reasoning from sentiment analysis of online product reviews. A major advantage is that the proposed method can extract extraordinary cases effectively with a large amount of online product reviews, based on which customer needs elicited with ordinary cases can be adapted to eliciting latent customer needs. This process does not need to interview a large number of lead users and thus is much more efficient in terms of time and cost.

Second, while traditional customer preference modeling and quantification methods do not consider subjective experiences, the proposed customer preference model based on cumulative prospect theory is able to accommodate cognitive tendency, affective influence, and risk attitudes. It helps customers make better product choices in terms of the holistic product utility. The hierarchical Bayesian model with MCMC is able to produce both group-level customer preference models and individual-level ones for customer homogeneity and individual differences, respectively.

Third, in order to make uses of social network effects, it is important to understand the product diffusion mechanism, especially product adoption prediction. We address several limitations of current diffusion models and propose a linear threshold-hurdle model, based on which a rough set technique is used to successfully predict product adoption. The hurdle utility is compared with the product holistic utility in the product adoption process. The results show the significant improvement by the incorporation of hurdle in the linear threshold-hurdle model.

Fourth, in order to evaluate viral product design, I formulate it as a game theoretic optimization problem, in which the leader makes use of both viral product attributes (such as those obtained by the method in Chapter 8) and viral influence attributes (i.e., seed customers) for adoption maximization, and the follower maximizes a shared surplus between customer satisfaction (in terms of the holistic product utility) and engineering cost for product line performance optimization. The optimal product configurations and top n seed customers obtained by a coordinate-wise solution strategy are expected to be the equilibrium solution of such a non-collaborative game between the leader and the follower.

9.2 Contributions

The major contributions of the dissertation manifest themselves through the proposal and development of a coherent framework of viral product design for social network effects. The deliverables are entailed in the strategy, fundamentals, methodology, validation, and application aspects, as elaborated below:

(1) At the strategy level, the following consensuses have been achieved (Chapters 1 and 2):

- Propose viral product design for social network effects;

- Highlight the importance of social network effects in terms of viral product attributes and viral influence attributes in the product adoption process.

(2) At the fundamental level, the following findings have been obtained (Chapter 3 and 4):

- Analyze the fundamentals of product design incorporating peer influence of social networks, including the identification of viral attributes, social network effects, relationships between the SoC problem and the InfMax problem, the comparison between the hurdle utility and the holistic product utility, customer preference modeling, diffusion mechanism, and marketing-engineering coordination, and so on ;
- Bridge the gaps between engineering design in terms of product portfolio planning and viral marketing in terms of product adoption maximization;
- Identify four design steps along the proposed technical framework, including latent customer needs elicitation, customer preference modeling and quantification, social network modeling, and viral product design evaluation.

(3) In terms of the methodology and supporting tools, the following deliverables have been promised (Chapters 5, 6, 7, and 8):

- Elicit latent customer needs by use case analogical reasoning from sentiment analysis of online product reviews;
- Examine subjective experiences in preference-based customers' product choice decision making using cumulative prospect theory;
- Model product adoption diffusion in large social network with a linear threshold-hurdle model, based on which product adoption is predicted with rough set theory;
- Joint maximize product adoption and shared surplus with a bi-level game theoretic optimization technique for viral product design evaluation.

(4) As for validation and application, several experimental and case studies have been conducted, including:

- A case study of Kindle Fire HD tablets to illustrate the process of latent customer needs elicitation (Chapter 5);

- An experiment used to test the influence of subjective experiences on customer preference-based product choice decision making (Chapter 6);
- A case study of Kindle Fire HD tablets to justify the advantages of the proposed linear threshold-hurdle model for product adoption prediction (Chapter 7);
- A case study of Kindle Fire HD tablets to evaluate viral product design for social network effects (Chapter 8).

9.3 Limitations

As an exploratory study of the proposed viral product design for social network effects, it suffers several limitations.

(1) The scope of viral attributes: In this dissertation, I extract a number of product attributes and attribute levels, which are considered as the candidates of viral product attributes. Therefore, the design space in terms of viral product attributes is limited by the product line itself. Although I extract new product attributes by satisfying latent customer needs elicited by use case analogical reasoning. However, it is possible that similar products to Kindle Fire HD tablets, such as iPad series, can have different product attributes from a completely new product line. In this sense, it is necessary to expand the space of viral product attributes by sentiment analysis of product reviews of similar products. Furthermore, other research suggests that a variety of product attributes could affect the degree of peer influence and social contagion in the product adoption process (Aral et al., 2013). For example, Berger and Milkman (2010) find that awe-inspiring news stories that are practically useful, surprising, positive, or affect-laden are more likely to make it into the New York Times "most e-mailed" articles list. Heath et al. (2001) show that disgusting urban legends are more likely to be shared. Hence, they suggest that affect-rich product attributes tend to be viral, which can be a fruitful direction for future research that connects emotional design (Zhou et al., 2010).

In terms of viral influence attributes, I only consider seed customers as one kind of viral influence attributes that can be manipulated based on their importance and influential ability. Many of the viral influence attributes are actually embedded in the social network websites, such as personalized referrals, automated broadcast notifications, and tagging, commenting on like or dislike, and inputting customer reviews or satisfaction levels. It seems that there is not much that I can manipulate like viral product attributes in order to maximize product awareness among social network users. However, firms can actually create social network platforms, such as product review sections of shopping websites. How they can make use of these viral influence attributes, even personalized incentive strategy for seed customers, to promote themselves, to communicate with their customers and the society, and to create new features and add-ons that can contribute to product adoption, can further complement the current research.

(2) The influence of product types on virality: In the domain of viral marketing, not all of the products seem to be able to take off in a social network. This is consistent with the idea proposed by Rogers (2003), who argues that four main elements influence the spread of a new idea: the innovation itself, communication channels, time, and a social system. This shows that the nature of a product, to some extent, determines the quality of being viral. It is suggested that the virality of technology-related products, especially in the digital form, tend to be magnified, such as referral links appended to e-mails sent from a Hotmail account and automated notifications of a user's activity sent by Facebook applications to the user's Facebook friends (Aral et al., 2013). In this research, only one product (line), i.e., Kindle Fire HD tablets, is used as a case study. This kind of product is somewhere between traditional physical products (e.g., an alarm clock) and digital technology-related products. Thus, the results may only be typical to this kind of products. Hence, it is necessary to include different types of products for experimentation.

(3) The influence of social network types on viral product design: In this dissertation, I construct the social network based on the reviewer-commenter links from Amazon.com. Its degree distribution on a log-log scale is similar to a power law, which implies that it is a scale-free network in terms of the structure (Onnela et al., 2007). Other types of social networks also exist, such as random networks (Bollobás, Random Graphs) and small world networks (Watts and Strogatz, 1998). These different network structures may influence the product adoption process to different extents. On the other hand, the social network is constructed based on online product reviews. This type of social network is essentially product-centric. However, other types of social networks, such as Flickr (a social website to share personal photos), Youtube (a social website to share videos), and LinkedIn (a business-oriented social networking service, mainly for professional networking) also specialize in a particular type of social communication. It is unknown how the network relationship formation and content on such platforms will influence viral product design. More general social networks with a wide variety of topics, such as Facebook and Twitter are also worth investigating. The former is a non-directional social network, while the latter is directional. It is also not clear that how the direction of a social network and how the noise content irrelevant to a product will influence viral product design.

9.4 Future Work

Social media and social networks have fundamentally changed the way we live. Several ideas are elaborated below for potential endeavors in the future.

(1) Traditional social networks vs. Online Social networks: Diffusion of innovations has been widely studied in social sciences with its focus on the spread and use of ideas from one social entity to another or one culture to another in the context of traditional social networks (Rogers, 2003). Social networks form a social system, in which how social entities interact determine the diffusion process to a large extent. By comparing traditional social networks with online social networks, I believe the differences in communication channels may influence the diffusion process to a large extent, such as speed and trust. Communications through social media are now one of the most rapid and efficient means of interactions among many people (Aral et al., 2013). On the other hand, interpersonal channels with a face-to-face exchange are more effective in persuading an individual to accept an innovation. Most individuals tend to evaluate an innovation through subjective assessment that is conveyed to them from someone who has already adopted the innovation (Aral et al., 2013). Prior empirical studies have shown that trust in information sources plays a major role in people's decision making (Renn and Levine, 1991). Thus, it is meaningful to study the factor of trust in the product adoption process, comparing traditional interpersonal social networks with online social networks.

(2) Mobile social media: Mobile social media combine mobile devices and social media so that user-generated content can be created and exchanged anywhere anytime with mobile devices (Kaplan, 2012). This has several impacts in terms of design and marketing a product. For example, while in the past customers had to print coupons, mobile social media allow tailored promotion to specific users with their mobile devices. For another instance, many mobile devices make use of location-aware technologies and time-sensitive fashions for retail location planning and product promotion (Ozimec et al., 2010). Others allow customers to scan 2-dimensional barcodes for transaction purposes. One forthcoming event is Apply Pay, which may digitalize credit cards in the future. Therefore, how a product should be designed to make use of mobile social media to maximize its adoption seems an interesting topic.

(3) Social media optimization: Social media optimization involves content generation using social media with the hopes of driving traffic and creating interest in a particular product, service, or topic (Frick, 2010). Many companies now integrate social media to manage their businesses, including knowledge management, brand building, customer satisfaction and relations, business development and more. By making use of social media optimization, it is important to combine human-computer interaction principles that can identify social technologies to promote certain customer behavior for various applications in healthcare, learning, education, and innovation. For example, Kamal et al. (2014) design an online social network, VivoSpace, in order to maximize health behavior change (e.g., being physically active to prevent chronic diseases) effectiveness by utilizing socially-based determinants in the design process. Another good example is massive open online courses (MOOCs) that aim at unlimited participation and open access via the web for individuals who want to learn (Pappano, 2014). Its success to some degree depends on how the platform is designed and optimized for social interactions (e.g., www.coursera.org, www.edx.org). In addition to traditional course materials, such as videos, readings, and problem sets, MOOCs provide an interactive community, which optimizes communications among students, professors, and teaching assistants without location restrictions. Hence, integrating these social media optimization techniques in the design process can be promising for some products and services.

APPENDIX A: DEFINITION OF PRECISION, RECALL, AND *F*-SCORE

We use F-score to measure the accuracy of the prediction/classification results. F-score is defined as the harmonic mean of precision and recall (Zhou et al., 2011b). Precision is a measure of exactness or fidelity, whereas recall is a measure of completeness. For a multi-class classification problem, assume there are N_c different classes, and the i-th class has a total number of N_i instances in the dataset. If the model predicts correctly C_i for the i-th class and predicts C_i^* instances to be in the i-th class, in which they actually belong to other classes, then the previous measures are defined as follows:

$$\text{Precision} = \frac{C_i}{C_i + C_i^*}, \tag{A1}$$

$$\text{Recall} = \frac{C_i}{N_i}, \tag{A2}$$

$$\text{F-score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}. \tag{A3}$$

APPENDIX B: IF-THEN RULES INVOLVED IN CASE-BASED REASONING

1) IF the user type is kids/students, THEN the need is conditioned with parental-control or is children appropriate.
2) IF the user type is the elderly, THEN the need is conditioned with an extremely easy interface.
3) IF the interaction environment is outdoor with sunlight, THEN the need is conditioned with outdoor with sunlight.
4) IF the interaction environment is dark indoor, THEN the need supports feasible operations in the dark environment.
5) IF the contextual event is a trip, THEN the need is conditioned with trip characteristics (e.g., no WIFI connection, no power available).
6) IF the contextual event is cooking, THEN the need is cooking friendly (e.g., it has a stand for supporting the tablet and looking up recipes).
7) IF the contextual event is working out, THEN the need supports activities related to working out (e.g., being fit for the gym equipment and offering applications with health related information).

REFERENCES

Aamodt, A., Plaza, E., 1994. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications, 7(1), 39-59.

Agrawal, R., Imieliński, T., Swami, A., 1993. Mining association rules between sets of items in large databases. Proceedings of the ACM SIGMOD International Conference on Management of Data, Washington, D.C., USA.

Ahn, H., 2010. Modeling and Analysis of Affective Influences on Human Experience, Prediction, Decision Making, and Behavior. PhD thesis, MIT.

Ahn, H., Picard, R.W., 2005. Affective-Cognitive Learning and Decision Making: A Motivational Reward Framework for Affective Agents, The 1st International Conference on Affective Computing and Intelligent Interaction (ACII 2005), Beijing, China.

Allenby, G.M., Rossi, P.E., 2006. Hierarchical Bayes Models: A Practitioner's Guide, in: Grover, R., Vriens, M. (Eds.), The Handbook of Marketing Research. Sage, Thousand Oaks, CA, pp. 418-440.

Alyaqout, S.F., Peters, D.L., Papalambros, P.Y., Ulsoy, A.G., 2011. Generalized Coupling Management in Complex Engineering Systems Optimization. ASME Journal of Mechanical Design, 133(9), 091005.

Aral, S., 2011. Commentary-Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion. Marketing Science, 30(2), 217-223.

Aral, S., Dellarocas, C., Godes, D., 2013. Introduction to the Special Issue-Social Media and Business Transformation: A Framework for Research. Information Systems Research, 24(1), 3-13.

Aral, S., Muchnik, L., Sundararajan, A., 2009. Distinguishing Influence Based Contagion from Homophily Driven Diffusion in Dynamic Networks. Proceedings of the National Academy of Sciences, 106(51), 21544–21549.

Aral, S., Walker, D., 2011. Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks. Management Science, 57(9), 1623-1639.

Archak, N., Ghose, A., Ipeirotis, P.G., 2007. Show Me the Money!: Deriving the Pricing Power of Product Features by Mining Consumer Reviews, Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, San Jose, California, USA, pp. 56-65.

Bampo, M., Ewing, M.T., Mather, D.R., Stewart, D., Wallace, M., 2008. The Effects of the Social Structure of Digital Networks on Viral Marketing Performance. Information Systems Research, 19(3), 273-290.

Banks, H.T., Grove, S., Hu, S., Ma, Y., 2005. A Hierarchical Bayesian Approach for Parameter Estimation in Hiv Models. Inverse Problems, 21(6), 1803-1822.

Barbieri, N., Bonchi, F., 2014. Influence Maximization with Viral Product Design, SIAM International Conference on Data Mining (SDM), Philadelphia, PA, USA.

Bard, J.F., 1998. Practical Bilevel Optimization: Algorithms and Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bazan, J.G., Szczuka, M., 2001. RSES and RSESLib - A Collection of Tools for Rough Set Computations, in: Ziarko, W., Yao, Y. (Eds.), RSCTS 2000, LNAI 2005. Springer-Verlag, Berlin/Heidelberg, pp. 106-113.

Bearden, W.O., Etzel, M.J., 1982. Reference Group Influence on Product and Brand Purchase Decisions. Journal of Consumer Research, 9(2), 183–194.

Bell, D.E., 1985. Disappointment in Decision-Making under Uncertainty. Operations Research, 33(1), 1-27.

Bello, J., Rolfe, M., 2014. Is Influence Mightier Than Selection? Forging Agreement in Political Discussion Networks During a Campaign. Social Networks, 36(0), 134-146.

Belloni, A., Freund, R., Selove, M., Simester, D., 2008. Optimizing Product Line Designs: Efficient Methods and Comparisons. Management Science, 54(9), 1544-1552.

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press, Cambridge, MA, USA.

Berger, J., Milkman, K., 2010. Social Transmission and Viral Culture. The Wharton School of the Univeristy of Pennsylvania, Philadelphia.

Bhagat, S., Goyal, A., Lakshmanan, L.V.S., 2012. Maximizing Product Adoption in Social Networks, Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, ACM, Seattle, Washington, USA, pp. 603-612.

Bhatt, R., Chaoji, V., Parekh, R., 2010. Predicting Product Adoption in Large-Scale Social Networks, Proceedings of the 19th ACM International Conference on Information and Knowledge Management. ACM, Toronto, ON, Canada, pp. 1039-1048.

Bickart, B., Schindler, R.M., 2001. Internet Forums as Influential Sources of Consumer Information. Journal of Interactive Marketing, 15(3), 31-40.

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G.A., Reynar, J., 2008. Building a Sentiment Summarizer for Local Service Reviews, WWW Workshop on NLP in the Information Explosion Era, Beijing, China.

Bollobás, B., Random Graphs. 2001, 2nd ed. Cambridge University Press, Cambridge, UK.

Bonacich, P., 2007. Some Unique Properties of Eigenvector Centrality. Social Networks, 29(4), 555-564.

Bouguila, N., Elguebaly, T., 2012. A Fully Bayesian Model Based on Reversible Jump Mcmc and Finite Beta Mixtures for Clustering. Expert Systems with Applications, 39(5), 5946-5959.

Bracha, A., Brown, D., 2012. Affective Decision Making: A Theory of Optimism Bias. Games and Economic Behavior, 75(1), 67-80.

Bradley, M.M., Lang, P.J., 1994. Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), 49-59.

Bradley, M.M., Lang, P.J., 1999. Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings, Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brands, T., van Berkum, E.C., 2014. Performance of a Genetic Algorithm for Solving the Multi-Objective, Multimodal Transportation Network Design Problem International Journal of Transportation, 2(1), 1-20.
Brandstätter, E., Gigerenzer, G., Hertwig, R., 2006. The Priority Heuristic: Making Choices without Trade-offs. Psychological Review, 113(2), 409-432.
Brown, A.S., 2008. The New Point of View: Focus on Design for Human Factors. ASME Mechanical Engineering Magazine, February Issue.
Budanitsky, A., Hirst, G., 2001. Semantic Distance in WordNet: An Experimental, Application-Oriented Evaluation of Five Measures, Workshop on WordNet and Other Lexical Resources, Second Meeting of the North American Chapter of the Association for Computational Linguistics, Pittsburgh, PA, pp. 29-34.
Burt, R.S., 1987. Social Contagion and Innovation: Cohesion Versus Structural Equivalence. American Journal of Sociology, 92(6), 1287-1335.
Calvete, H.I., Galé, C., Mateo, P.M., 2008. A New Approach for Solving Linear Bilevel Problems Using Genetic Algorithms. European Journal of Operational Research, 188(1), 14-28.
Camm, J.D., Cochran, J.J., Curry, D.J., Kannan, S., 2006. Conjoint Optimization: An Exact Branch-and-Bound Algorithm for the Share-of-Choice Problem. Management Science, 52(3), 435-447.
Carenini, G., Ng, R.T., Zwart, E., 2005. Extracting Knowledge from Evaluative Text, Proceedings of the 3rd International Conference on Knowledge Capture. ACM, Banff, Alberta, Canada, pp. 11-18.
Carlgren, L., 2013. Identifying Latent Needs: Towards a Competence Perspective on Attractive Quality Creation. Total Quality Management & Business Excellence, 24(11-12), 1347-1363.
Centola, D., 2010. The Spread of Behavior in an Online Social Network Experiment. Science, 329(5996), 1194-1197.
Chen, L., Qi, L., Wang, F., 2012. Comparison of Feature-Level Learning Methods for Mining Online Consumer Reviews. Expert Systems with Applications, 39(10), 9588-9601.
Chen, P.-Y., Dhanasobhon, S., Smith, M.D., 2008. All Reviews Are Not Created Equal: The Disaggregate Impact of Reviews and Reviewers at Amazon.com. Available at SSRN: http://ssrn.com/abstract=918083.
Chen, W., Collins, A., Cummings, R., Ke, T., Liu, Z., Rincón, D., Sun, X., Wang, Y., Wei, W., Yuan, Y., 2011a. Influence Maximization in Social Networks When Negative Opinions May Emerge and Propagate. Proceedings of SDM, 379-390.
Chen, W., Hoyle, C., Wassenaar, H., 2013a. A Choice Modeling Approach for Usage Context-Based Design, Decision-Based Design. Springer, London, pp. 255-285.
Chen, W., Hoyle, C., Wassenaar, H., 2013b. Decision-Based Design Framework, Decision-Based Design. Springer, London, pp. 79-105.
Chen, W., Hoyle, C., Wassenaar, H., 2013c. Decision-Based Design: An Approach for Enterprise-Driven Engineering Design, Decision-Based Design. Springer, London, pp. 3-11.
Chen, W., Hoyle, C., Wassenaar, H., 2013d. Fundamentals of Analytical Techniques for Modeling Consumer Preferences and Choices, Decision-Based Design. Springer, London, pp. 35-77.

Chen, W., Hoyle, C., Wassenaar, H.J., 2013e. Decision-Based Design Integrating Consumer Preferences into Engineering Design. Springer, London.

Chen, W., Wang, C., Wang, Y., 2010. Scalable Influence Maximization for Prevalent Viral Marketing in Large-Scale Social Networks, Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, Washington, DC, USA, pp. 1029-1038.

Chen, W., Wang, Y., Yang, S., 2009. Efficient Influence Maximization in Social Networks, 15th ACM SIGKDD International Conference Knowledge Discovery and Data Mining, Paris, France, pp. 199–208.

Chen, Y.-D., Brown, S.A., Hu, P.J.-H., King, C.-C., Chen, H., 2011b. Managing Emerging Infectious Diseases with Information Systems: Reconceptualizing Outbreak Management through the Lens of Loose Coupling. Information Systems Research, 22(3), 447-468.

Chen, Y., Xie, J., 2008. Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. Management Science, 54(3), 477-491.

Cho, C.K., Kim, Y.S., Lee, W.J., 2010. Economical, Ecological and Experience Values for Product-Service Systems, The 7th Design & Emotion Conference, Chicago, USA.

Chow, J.Y.J., Lee, G., Yang, I., 2010. Genetic Algorithm to Estimate Cumulative Prospect Theory Parameters for Selection of High-Occupancy-Vehicle Lane. Transportation Research Record: Journal of the Transportation Research Board, 2(2157), 71-77.

Clauset, A., Newman, M.E.J., Moore, C., 2004. Finding Community Structure in Very Large Networks. Physical Review E, 70(6), 066111.

Clemons, E., Gao, G., Hitt, L., 2006. When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry. Journal of Management Information Systems, 23(2), 149-171.

Colson, B., Marcotte, P., Savard, G., 2005. Bilevel Programming: A Survey. 4OR, 3(2), 87-107.

Colson, B., Marcotte, P., Savard, G., 2007. An Overview of Bilevel Optimization. Annals of Operations Research, 153(1), 235-256.

Cowles, M.K., Carlin, B.P., 1996. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. Journal of the American Statistical Association, 91(434), 883-904.

Crandall, B., Klein, G., Hoffman, R., 2006. Working Minds: A Practitioner's Guide to Cognitive Task Analysis. The MIT Press, Cambridge, MA, USA.

Darren, Q., Liming, C., Maurice, M., 2012. Social Network Analysis: A Survey. International Journal of Ambient Computing and Intelligence, 4(3), 46-58.

Deb, K., Sinha, A., 2010. An Efficient and Accurate Solution Methodology for Bilevel Multi-Objective Programming Problems Using a Hybrid Evolutionary-Local-Search Algorithm. Evolutionary Computation, 18(3), 403-449.

Decker, R., Trusov, M., 2010. Estimating Aggregate Consumer Preferences from Online Product Reviews. International Journal of Research in Marketing, 27(4), 293-307.

Desvousges, W.H., Johnson, F., Dunford, R., Hudson, S., Wilson, K., Boyle, K., 1993. Measuring Resource Damages with Contingent Valuation: Tests of Validity and

Reliability, in: Hausman, J.A. (Ed.), Contingent Valuation: A Critical Assessment. North Holland, Amsterdam, pp. 91-164.
Devendorf, E., Lewis, K., 2011. The Impact of Process Architecture on Equilibrium Stability in Distributed Design. ASME Journal of Mechanical Design, 133(10), 101001-(101001-101011).
Dhar, R., Wertenbroch, K., 2000. Consumer Choice between Hedonic and Utilitarian Goods. Journal of Marketing Research, 37(1), 60-71.
Dichter, E., 1966. How Word-of-Mouth Advertising Works. Harvard Business Review, 44(6), 147-166.
Ding, X., Liu, B., Yu, P.S., 2008. A Holistic Lexicon-Based Approach to Opinion Mining, Proceedings of the 2008 International Conference on Web Search and Data Mining. ACM, Palo Alto, California, USA, pp. 231-240.
Domingos, P., Richardson, M., 2001. Mining the Network Value of Customers, 7th ACM SIGKDD International Conference Knowledge Discovery and Data Mining. ACM, San Francisco, CA, USA, pp. 57–66.
Dou, Y., Niculescu, M.F., Wu, D.J., 2013. Engineering Optimal Network Effects Via Social Media Features and Seeding in Markets for Digital Goods and Services. Information Systems Research, 24(1), 164-185.
Du, G., Jiao, R.J., Chen, M., 2014. Joint Optimization of Product Family Configuration and Scaling Design by Stackelberg Game. European Journal of Operational Research, 232(2), 330-341.
Ellsworth, P.C., Scherer, K.R., 2003. Appraisal Processes in Emotion, in: Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (Eds.), Handbook of Affective Sciences. Oxford University Press, New York, pp. 572-595.
Falcioni, J.G., 2008. Make It Work. ASME Mechanical Engineering Magazine, February Issue.
Fang, X., Hu, P.J.-H., Li, Z., Tsai, W., 2013. Predicting Adoption Probabilities in Social Networks. Information Systems Research, 24(1), 128-145.
Fisher, J.C., Pry, R.H., 1971. A Simple Substitution Model of Technological Change. Technological Forecasting & Social Change, 3(1), 75-88.
Frick, T., 2010. Social Media Optimization, in: Frick, T. (Ed.), Return on Engagement. Focal Press, Boston, MA, USA, pp. 205-228.
Friedkin, N.E., 1998. A Structural Theory of Social Influence. Cambridge University Press, Cambridge, UK.
Fruchterman, T.M.J., Reingold, E.M., 1991. Graph Drawing by Force-Directed Placement. Software: Practice and Experience, 21(11), 1129-1164.
Gambetti, E., Giusberti, F., 2012. The Effect of Anger and Anxiety Traits on Investment Decisions. Journal of Economic Psychology, 33(6), 1059-1069.
Garrow, L.A., Koppelman, F.S., 2004. Multinomial and Nested Logit Models of Airline Passengers' No-Show and Standby Behaviour. Journal of Revenue Pricing and Management, 3(3), 237-253.
Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge, UK.
Geman, S., Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(6), 609-628.

Genat, R. (2004). The American Car Dealership. MBI Publishing Company LLC, St. Paul, MN, USA.
Ghani, R., Probst, K., Liu, Y., Krema, M., Fano, A., 2006. Text Mining for Product Attribute Extraction. ACM SIGKDD Explorations Newsletter, 8(1), 41-48.
Ghose, A., Ipeirotis, P.G., 2011. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. IEEE Transactions on Knowledge and Data Engineering, 23(10), 1498-1512.
Goel, A., Craw, S., 2005. Design, Innovation and Case-Based Reasoning. Knowledge Engineering Review 20(3): 271-276.
Goldenberg, J., Han, S., Lehmann, D.R., Hong, J.W., 2009. The Role of Hubs in the Adoption Process. Journal of Marketing, 73(2), 1-13.
Gora, G., Wojna, A., 2002. Riona: A New Classification System Combining Rule Induction and Instance-Based Learning. Fundamenta Informaticae, 51(4), 369-390.
Goyal, A., Bonchi, F., Lakshmanan, L.V.S., 2010. Learning Influence Probabilities in Social Networks, Proceedings of the Third ACM International Conference on Web Search and Data Mining. ACM, New York, USA, pp. 241-250.
Goyal, A., Lu, W., Lakshmanan, L.V.S., 2011. CELF++: Optimizing the Greedy Algorithm for Influence Maximization in Social Networks, Proceedings of the 20th International Conference Companion on World Wide Web. ACM, Hyderabad, India, pp. 47-48.
Granovetter, M., 1973. The Strength of Weak Ties. American Journal of Sociology, 78(May), 1360-1380.
Granovetter, M., 1978. Threshold Models of Collective Behavior. American Journal of Sociology, 83(May), 1420-1443.
Granovetter, M., Soong, R., 1988. Threshold Models of Diversity: Chinese Restaurants, Residential Segregation and the Spiral of Silence, in: Clogg, C. (Ed.), Sociological Methodology, pp. 69-104.
Gratch, J., Marsella, S., 2004. A Domain-Independent Framework for Modeling Emotion. Cognitive Systems Research, 5(4), 269-306.
Green, P.E., Krieger, A.M., 1985. Models and Heuristics for Product Line Selection. Marketing Science, 4(1), 1-19.
Green, P.E., Srinivasan, V., 1978. Conjoint Analysis in Consumer Behavior: Status and Outlook. Journal of Consumer Research, 5(Sept.), 103-123.
Green, P.E., Srinivasan, V., 1990. Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. Journal of Marketing, 54(4), 3-19.
Grover, V., Kettinger, W.J., 2000. Process Think: Winning Perspectives for Business Change in the Information Age. Idea Group Publisher, Hershey, PA, USA.
Gundecha, P., Liu, H., 2012. Mining Social Media: A Brief Introduction. Tutorials in Operations Research, 9(4), 1-17.
Gunnec, D., 2012. Integrating Social Network Effects in Product Design and Diffusion, Business and Management: Decision & Information Technologies. PhD Thesis, University of Maryland, Maryland.
Hand, D.J., Mannila, H., Smyth, P., 2001. Principles of Data Mining. The MIT Press, Cambridge, MA, USA.
Hannukainen, P., Hölttä-Otto, K., 2006. Identifying Customer Needs: Disabled Persons as Lead Users, ASME 2006 International Design Engineering Technical

Conferences and Computers and Information in Engineering Conference, Philadelphia, Pennsylvania, USA.
Hanski, J., Reunanen, M., Kunttu, S., Karppi, E., Lintala, M., Nieminen, H., 2014. Customer Observation as a Source of Latent Customer Needs and Radical New Ideas for Product-Service Systems, in: Lee, J., Ni, J., Sarangapani, J., Mathew, J. (Eds.), Engineering Asset Management 2011. Springer, London, pp. 395-407.
Harrison, G.W., Rutström, E.E., 2009. Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral. Experimental Economics, 12(2), 133-158.
Hartmann, W.R., 2010. Demand Estimation with Social Interactions and the Implications for Targeted Marketing. Marketing Science, 29(4), 585-601.
Hassenzahl, M., 2004. The Interplay of Beauty, Goodness and Usability in Interactive Products. Human Computer Interaction, 19(4), 319-349.
Hauser, J.R., Shugan, S.M., 1980. Intensity Measures of Consumer Preferences. Operations Research, 28(2), 278-320.
Hausman, J., McFadden, D., 1984. Specification Tests for the Multinomial Logit Model. Econometrica, 52(5), 1219-1240.
Hayes, C., Goel, A., Tumer, I., Agogino, A., Regli, W., 2011. Intelligent Support for Product Design: Looking Backwards, Looking Forwards. ASME Journal of Computing and Information Science in Engineering, 11(2), 021007.
Hazelrigg, G.A., 1998. A Framework for Decision-Based Engineering Design. Journal of Mechanical Design, 120(4), 653-658.
Heath, C., Bell, C., Sternberg, E., 2001. Emotional Selection in Memes:The Case of Urban Legends. Journal of Personality and Social Psychology, 81(6), 1028–1041.
Ho, T.H., Tang, C.S., 1998. Product Variety Management: Research Advances. Kluwer Academic Publishers, Boston, MA, USA.
Holland, J.H., 1992. Adaptation in Natural and Artificial Systems. The MIT Press, Cambridge, MA, USA.
Howard, T., 2005. 'Viral' Advertising Spreads through Marketing Plans, USA Today.
Hsee, C.K., Rottenstreich, Y., 2004. Music, Pandas, and Muggers: On the Affective Psychology of Value. Journal of Experimental Psychology: General, 133(1), 23-30.
Hu, M., Liu, B., 2004a. Mining and Summarizing Customer Reviews, Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, Seattle, WA, USA, pp. 168-177.
Hu, M., Liu, B., 2004b. Mining Opinion Features in Customer Reviews, The 19th National Conference on Artificial Intellgience, San Jose, CA, USA.
Isen, A.M., 2001. An Influence of Positive Affect on Decision Making in Complex Situations: Theoretical Issues with Practical Implications. Journal of Consumer Psychology, 11(2), 75-85.
Iyengar, R., van den Bulte, C., Valente, T.W., 2011. Opinion Leadership and Social Contagion in New Product Diffusion. Marketing Science, 30(2), 195-212.
Jeroslow, R., 1985. The Polynomial Hierarchy and a Simple Model for Competitive Analysis. Mathematical Programming, 32(2), 146-164.
Ji, Y., Jiao, R.J., Chen, L., Wu, C., 2013. Green Modular Design for Material Efficiency: A Leader-Follower Joint Optimization Model Based on Constrained Genetic Algorithm. Journal of Cleaner Production, 41(2), 187-201.

Jiao, J., Chen, C.-H., 2006. Customer Requirement Management in Product Development: A Review of Research Issues. Concurrent Engineering: Research and Applications, 14(3), 173-185.

Jiao, J., Tseng, M.M., 1999. An Information Modeling Framework for Product Families to Support Mass Customization Production. CIRP Annals, 48(1), 93-98.

Jiao, J., Tseng, M.M., 2004. Customizability Analysis in Design for Mass Customization. Computer-Aided Design, 36(8), 745-757.

Jiao, J., Xu, Q., Du, J., Zhang, Y., Helander, M.G., Khalid, H.M., Helo, P., Ni, C., 2007a. Analytical Affective Design with Ambient Intelligence for Mass Customization and Personalization. International Journal of Flexible Manufacturing Systems, 19(4), 570-595.

Jiao, J., Zhang, Y., 2005a. Product Portfolio Identification Based on Association Rule Mining. Computer-Aided Design, 37(2), 149-172.

Jiao, J., Zhang, Y., 2005b. Product Portfolio Planning with Customer-Engineering Interaction. IIE Transactions, 37(9), 801-814.

Jiao, J., Zhang, Y., Wang, Y., 2007b. A Heuristic Genetic Algorithm for Product Portfolio Planning. Computers & Operations Research, 34(6), 1777-1799.

Jin, W., Ho, H.H., Srihari, R.K., 2009. Opinionminer: A Novel Machine Learning System for Web Opinion Mining and Extraction, Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Paris, France, pp. 1195-1204.

Kahneman, D., 2000. Experienced Utility and Objective Happiness: A Moment-Based Approach, in: Kahneman, D., Tversky, A. (Eds.), Choices, Values, and Frames. Cambridge University Press, New York, pp. 673-692.

Kahneman, D., 2003. A Perspective on Judgment and Choice: Mapping Bounded Rationality. American Psychologist, 58(9), 697-720.

Kahneman, D., Ritov, I., Schkade, D., 1999. Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues. Journal of Risk and Uncertainty, 19(1-3), 203-235.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 1251-1289.

Kalish, S., 1985. A New Product Adoption Model with Price, Advertising, and Uncertainty. Management Science, 31(12), 1569-1585.

Kamal, N., Fels, S., Fergusson, M., 2014. Online Social Networks for Health Behaviour Change: Designing to Increase Socialization. Computers in Human Behavior, In Press, http://dx.doi.org/10.1016/j.chb.2014.03.068.

Kano, N., Seraku, N., Takahashi, F., Tsuji, S., 1984. Attractive Quality and Must-Be Quality. Journal of the Japanese Society for Quality Control, 14(2), 39-48.

Kaplan, A.M., 2012. If You Love Something, Let It Go Mobile: Mobile Marketing and Mobile Social Media 4x4. Business Horizons, 55(2), 129-139.

Kelman, H.C., 1961. Processes of Opinion Change. Public Opinion Quarterly, 25(1), 57.

Kempe, D., Kleinberg, J., Tardos, É., 2003. Maximizing the Spread of Influence through a Social Network, Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Washington, D.C., USA, pp. 137-146.

Khan, U., Dhar, R., Wertenbroch, K., 2004. A Behavioral Decision Theoretic Perspective on Hedonic and Utilitarian Choice. www.insead.edu/facultyresearch/research/doc.cfm?did=1411.
Kimura, M., Saito, K., 2006. Tractable Models for Information Diffusion in Social Networks, Proceedings of the 10th European conference on Principle and Practice of Knowledge Discovery in Databases. Springer-Verlag, Berlin, Germany, pp. 259-271.
Kinnear, T.C., Taylor, J.R., 1995. Marketing Research: An Applied Approach, 5th ed. McGraw-Hill, New York.
Kleinberg, J., 2008. The Convergence of Social and Technological Networks. Communication of the ACM, 51(11), 66–72.
Kohli, R., Krishnamurti, R., 1989. Optimal Product Design Using Conjoint Analysis: Computational Complexity and Algorithms. European Journal of Operational Research, 40(2), 186-195.
Kokkolaras, M., Mourelatos, Z.P., Papalambros, P.Y., 2006. Design Optimization of Hierarchically Decomposed Multilevel System under Uncertainty. ASME Journal of Mechanical Design, 128(2), 503-508.
Kumar, V., Bhaskaran, V., Mirchandani, R., Shah, M., 2013. Creating a Measurable Social Media Marketing Strategy: Increasing the Value and Roi of Intangibles and Tangibles for Hokey Pokey. Marketing Science, 32(2), 194-212.
Lai, Y.-J., 1996. Hierarchical Optimization: A Satisfactory Solution. Fuzzy Sets and Systems, 77(3), 321-335.
Law, E.L.C., Van Schaik, P., 2010. Modelling User Experience - an Agenda for Research and Practice. Interacting with computers, 22(5), 313-322.
Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Van Alstyne, M., 2009. Computational Social Science. Science, 323(5915), 721-723.
Lee, M.D., Newell, B.R., 2011. Using Hierarchical Bayesian Methods to Examine the Tools of Decision-Making. Judgment and Decision Making, 6(8), 832-842.
Lee, T.Y., 2007. Needs-Based Analysis of Online Customer Reviews, Proceedings of the ninth international conference on Electronic commerce. ACM, Minneapolis, MN, USA, pp. 311-318.
Lerner, J.S., Keltner, D., 2001. Fear, Anger, and Risk. Journal of Personality and Social Psychology, 81(1), 146-159.
Leskovec, J., Horvitz, E., 2008. Planetary-Scale Views on a Large Instant-Messaging Network, Proceedings of the 17th International Conference on World Wide Web. ACM, Beijing, China, pp. 915-924.
Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., Glance, N., 2007. Cost-Effective Outbreak Detection in Networks, The 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, San Jose, CA, USA, pp. 420-429.
Leung, Y.-W., Wang, Y., 2001. An Orthogonal Genetic Algorithm with Quantization for Global Numerical Optimization. IEEE Transactions on Evolutionary Computation, 5(1), 41-53.
Levy, D.A., 1992. The Liberating Effects of Interpersonal Influence: An Empirical Investigation of Disinhibitory Contagion. Journal of Social Psychology, 132(4), 469–473.

Lewis, K., Chen, W., Schmidt, L., 2006. Decision Making in Engineering Design. ASME Press, New York.

Li, H., Azarm, S., 2002. An Approach for Product Line Design Selection under Uncertainty and Competition. Journal of Mechanical Design, 124(3), 385-392.

Li, H., Zhang, L., Jiao, Y., 2014. Solution for Integer Linear Bilevel Programming Problems Using Orthogonal Genetic Algorithm. Journal of Systems Engineering and Electronics, 25(3), 443-451.

Lin, C.-F., Wang, S.-D., 2002. Fuzzy Support Vector Machines. IEEE Transactions on Neural Networks and Learning Systems, 13(2), 464-471.

Lin, J., Seepersad, C.C., 2007. Empathic Lead Users: The Effects of Extraordinary User Experiences on Customer Needs Analysis and Product Redesign, ASME International Design Engineering Technical Conferences/Computers and Information in Engineering Conference, Las Vegas, NV, USA, pp. 289-296.

Liu, B., 2010. Sentiment Analysis and Subjectivity, in: Indurkhya, N., Damerau, F.J. (Eds.), Handbook of Natural Lanuage Processing, 2nd ed. Chapman and Hall/CRC, Boca Raton, FL, USA.

Liu, B., Zhang, L., 2012. A Survey of Opinion Mining and Sentiment Analysis, in: Aggarwal, C.C., Zhai, C. (Eds.), Mining Text Data. Springer US, pp. 415-463.

Liu, Y., Jin, J., Ji, P., Harding, J.A., Fung, R.Y.K., 2013. Identifying Helpful Online Reviews: A Product Designer's Perspective. Computer-Aided Design, 45(2), 180-194.

Louviere, J.J., 1988. Analyzing Decision Making: Metric Conjoint Analysis. Sage, London.

Louviere, J.J., Flynn, T.N., Carson, R.T., 2010. Discrete Choice Experiments Are Not Conjoint Analysis. Journal of Choice Modelling, 3(3), 57-72.

Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated Choice Methods: Analysis and Applications. Cambridge University Press, New York.

Lu, W., Lakshmanan, L.V.S., 2012. Profit Maximization over Social Networks, Proceedings of the 2012 IEEE 12th International Conference on Data Mining. IEEE Computer Society, pp. 479-488.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. Winbugs - a Bayesian Modelling Framework: Concepts, Structure, and Extensibility. Statistics and Computing, 10(4), 325-337.

Luo, L., 2011. Product Line Design for Consumer Durables: An Integrated Marketing and Engineering Approach. Journal of Marketing Research, 48(1), 128-139.

Luo, Z.Q., Pang, J.S., Ralph, D., 1996. Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, UK.

Malmkjær, K., 2002. The Linguistics Encyclopedia, 2nd ed. Routledge, New York.

Marinier Iii, R.P., Laird, J.E., Lewis, R.L., 2009. A Computational Unification of Cognitive Behavior and Emotion. Cognitive Systems Research, 10(1), 48-69.

McFadden, D., Train, K., 2000. Mixed Mnl Models for Discrete Response. Journal of Applied Econometrics, 15(5), 447-470.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a Feather: Homophily in Social Networks. Annual Review of Sociology, 27, 415-444.

Mellers, B.A., 2000. Choice and Relative Pleasure of Consequences. Psychological Bulletin, 126(6), 910-924.

Meth, H., Brhel, M., Maedche, A., 2013. The State of the Art in Automated Requirements Elicitation. Information and Software Technology, 55(10), 1695-1709.
Miao, Q., Li, Q., Dai, R., 2009. Amazing: A Sentiment Mining and Retrieval System. Expert Systems with Applications, 36(3), 7192-7198.
Michalek, J.J., Ebbes, P., Adigüzel, F., Feinberg, F.M., Papalambros, P.Y., 2011. Enhancing Marketing with Engineering: Optimal Product Line Design for Heterogeneous Markets. International Journal of Research in Marketing, 28(1), 1-12.
Miller, G.A., 1995. Wordnet: A Lexical Database for English. Communications of the ACM, 38(11), 39-41.
Nair, S.K., Thakur, L.S., Wen, K., 1995. Near Optimal Solutions for Product Line Design and Selection: Beam Search Heuristics. Management Science, 41(5), 767-785.
Narayan, V., Rao, V.R., Saunders, C., 2011. How Peer Influence Affects Attribute Preferences: A Bayesin Updating Mechanism. Marketing Science, 30(2), 368-384.
Neal, R.M., 1993. Probabilistic Inference Using Markov Chain Monte Carlo Methods. Technical Report CRG-TR-93-1, University of Toronto, Toronto.
Newman, M.E.J., 2005. A Measure of Betweenness Centrality Based on Random Walks. Social Networks, 27(1), 39-54.
Newman, M.E.J., 2010. Networks: An Introduction. Oxford University Press, Oxford, UK.
Nilsson, H., Rieskamp, J., Wagenmakers, E.J., 2011. Hierarchical Bayesian Parameter Estimation for Cumulative Prospect Theory. Journal of Mathematical Psychology, 55(1), 84-93.
North, E., de Vos, R., 2002. The Use of Conjoint Analysis to Determine Consumer Buying Preferences: A Literature Review Journal of Family Ecology and Consumer Sciences, 30, 32-39.
Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., Barabási, A.-L., 2007. Structure and Tie Strengths in Mobile Communication Networks. Proceedings of the National Academy of Sciences, 104(18), 7332-7336.
Orsborn, S., Cagan, J., Boatwright, P., 2009. Quantifying Aesthetic Form Preference in a Utility Function. ASME Journal of Mechanical Design, 131(6), 061001-0610010.
Otto, K., Wood, K., 2001. Product Design: Techniques in Reverse Engineering and New Product Development. Prentice Hall, Upper Saddle River, NJ, USA.
Ozimec, A.-M., Natter, M., Reutterer, T., 2010. Geographical Information Systems-Based Marketing Decisions: Effects of Alternative Visualizations on Decision Quality. Journal of Marketing, 74(6), 94-110.
Pahl, G., Beitz, W., Feldhusen, J., Grote, K.H., 2007. Engineering Design a Systematic Approach, 3rd ed. Springer, London.
Pal, S.K., Shiu, S., 2004. Foundations of Soft Case-Based Reasoning. Wiley-Interscience, Hoboken, New Jersey, USA.
Pappano, L., 2014. The Year of the MOOC. The New York Times.
Park, J., Han, S.H., Kim, H.K., Oh, S., Moon, H., 2013. Modeling User Experience: A Case Study on a Mobile Device. International Journal of Industrial Ergonomics, 43(2), 187-196.
Park, Y., Lee, S., 2011. How to Design and Utilize Online Customer Center to Support New Product Concept Generation. Expert Systems with Applications, 38(8), 10638-10647.

Pawlak, Z., 1991. Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer, Dordrecht, The Netherlands.

Penolazzi, B., Gremigni, P., Russo, P.M., 2012. Impulsivity and Reward Sensitivity Differentially Influence Affective and Deliberative Risky Decision Making. Personality and Individual Differences, 53(5), 655-659.

Picard, R.W., 1997. Affective Computing. The MIT Press, Cambridge, MA, USA.

Picard, R.W., Vyzas, E., Healey, J., 2001. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10), 1175-1191.

Power, T.E., Swartzman, L.C., Robinson, J.W., 2011. Cognitive-Emotional Decision Making(Cedm): A Framework of Patient Medical Decision Making. Patient Education and Counseling, 83(2), 163-169.

Prasad, K.G.D., Subbaiah, K.V., Rao, K.N., Sastry, C.V.R.S., 2010. Prioritization of Customer Needs in House of Quality Using Conjoint Analysis. International Journal for Quality research, 4(2), 145-154.

Putthividhya, D., Hu, J., 2011. Bootstrapped Named Entity Recognition for Product Attribute Extraction, Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Edinburgh, UK, pp. 1557-1567.

Rai, R., 2012. Identifying Key Product Attributes and Their Importance Levels from Online Customer Reviews, ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago, IL, USA.

Raju, S., Pingali, P., Varma, V., 2009. An Unsupervised Approach to Product Attribute Extraction, Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval. Springer-Verlag, Toulouse, France, pp. 796-800.

Rencher, A.C., 1992. Interpretation of Canonical Discriminant Functions, Canonical Variates, and Principal Components. The American Statistician, 46(3), 217-225.

Renn, O., Levine, D., 1991. Credibility and Trust in Risk Communication, in: Kasperson, R., Stallen, P. (Eds.), Communicating Risks to the Public. Springer, Netherlands, pp. 175-217.

Rice, R.E., 1994. Using Network Concepts to Clarify Sources and Mechanisms of Social Influence, in: Richardss, W.D., Barn, G.A. (Eds.), Progress in Communication Sciences, Volume XII. Ablex, Norwood, NJ, USA.

Rice, R.E., Grand, A.E., Schmitz, J., Torobin, J., 1990. Individual and Network Influences on the Adoption and Perceived Outcomes of Electronic Messaging. Social Networks, 12(1), 27-55.

Richardson, M., Domingos, P., 2002. Mining Knowledge-Sharing Sites for Viral Marketing, The Eighth International Conference on Knowledge Discovery and Data Mining. ACM Press, Edmonton, Canada.

Rieder, B., 2013. Studying Facebook Via Data Extraction: The Netvizz Application, 5th Annual ACM Web Science Conference, Paris, France.

Riesbeck, C., Schank, R., 1989. Inside Case-Based Reasoning. Lawrence-Erlbaum, Hillsdale, NJ, USA.

Rogers, E.M., 2003. Diffusion of Innovations, 5th ed. Free Press, New York.

Romero, D.M., Galuba, W., Asur, S., Huberman, B.A., 2011. Influence and Passivity in Social Media, Proceedings of the 20th international Conference Companion on World Wide Web. ACM, Hyderabad, India, pp. 113-114.

Rook, D.W., 1987. The Buying Impulse. Journal of Consumer Research, 14(2), 189-199.

Rottenstreich, Y., Hsee, C.K., 2001. Money, Kisses, and Electric Shocks: On the Affective Psychology of Risk. Psychological Science, 12(3), 185-190.

Rouder, J.N., Lu, J., 2005. An Introduction to Bayesian Hierarchical Models with an Application in the Theory of Signal Detection. Psychonomic Bulletin & Review, 12(4), 573-604.

Scherer, K.R., Shorr, A., Johnstone, T., 2001. Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, Canary, NC, USA.

Seligman, M.E.P., 1975. Helplessness: On Depression, Development, and Death. W.H. Freeman, CA, USA.

Shao, Y., Lunettab, R.S., 2012. Comparison of Support Vector Machine, Neural Network, and Cart Algorithms for the Land-Cover Classification Using Limited Training Data Points. ISPRS Journal of Photogrammetry and Remote Sensing, 70(June), 18-87.

Shapiro, C., Varian, H.R., 1999. Information Rules: A Strategic Guide to the Network Economy. Harvard Business Review Press, MA, USA.

Simon, H., 1982. Affect and Cognition: Comments, in: Clark, M.S., Fiske, S.T. (Eds.), Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition. Erlbaum, NJ, USA, pp. 333-342.

Simon, H.A., 1956. Rational Choice and the Structure of the Environment. Psychological Review, 63(2), 129-138.

Simpson, T.W., 2004. Product Platform Design and Customization: Status and Promise. Artificial Intelligence for Engineering Design Analysis and Manufacturing, 18(1), 3-20.

Sinha, R., Parsons, O., 1996. Multivariate Response Patterning of Fear and Anger. Cognition and Emotion, 10(2), 173-198.

Slovic, P., Finucane, M.L., Peters, E., Macgregor, D.G., 2004. Risk as Analysis and Risk as Feelings: Some Thoughts About Affect, Reason, Risk, and Rationality. Risk Analysis, 24(2), 311-322.

Stackelberg, H., 1952. The Theory of Market Economy. Oxford University Press, Oxford, UK.

Stephen, A.T., Berger, J.A., 2009. Creating Contagious: How Social Networks and Item Characteristics Combine to Spur Ongoing Consumption and Reinforce Social Epidemics. Working paper, Wharton School, University of Pennsylvania, Philadelphia, PA, USA.

Steyvers, M., Lee, M.D., Wagenmakers, E.J., 2009. A Bayesian Analysis of Human Decision-Making on Bandit Problems. Journal of Mathematical Psychology, 53(3), 168-179.

Taguchi, G., 1995. Quality Engineering (Taguchi Methods) for the Development of Electronic Circuit Technology. IEEE Transactions on Reliability, 44(2), 225-229.

Taguchi, G., Chowdhury, S., Taguchi, S., 2000. Robust Engineering--Learn How to Boost Quality While Reducing Costs and Time to Market. McGraw-Hill, New York.

Thoai, N.V., Yamamoto, Y., Yoshise, A., 2002. Global Optimization Method for Solving Mathematical Programs with Linear Complementarity Constraints. Institute of Policy and Planning Sciences, University of Tsukuba, Japan.

Thurston, D.L., 2006. Utility Function Fundamentals, in: Lewis, K.E., W.Chen, Schmidt, L.C. (Eds.), Decision Making in Engineering Design. ASME Press, New York, pp. 15-19.

Thurston, D.L., Liu, T., 1991. Design Evaluation of Multiple Attributes under Uncertainty. International Journal of Control, Automation and Systems, 1(2), 143-159.

Titov, I., Mcdonald, R., 2008. A Joint Model of Text and Aspect Ratings for Sentiment Summarization, The 46th Annunal Meeting of Association for Computational Linguistics: Human Language Technologies, , Columbus, Ohio, USA.

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

Trusov, M., Bodapati, A.V., Bucklin, R.E., 2010. Determining Influential Users in Internet Social Networks. Journal of Marketing Research, 47(4), 643–658.

Tsai, J.-T., Liu, T.-K., Chou, J.-H., 2004. Hybrid Taguchi-Genetic Algorithm for Global Numerical Optimization. IEEE Transactions on Evolutionary Computation, 8(4), 365-377.

Tseng, M.M., Jiao, J., 1998. Computer-Aided Requirement Management for Product Definition: A Methodology and Implementation Concurrent Engineering: Research and Applications, 6(3), 145-160.

Tsvetovat, M., Kouznetsov, A., 2011. Social Network Analysis for Startups. O'Reilly Media, Sebastopol, CA, USA.

Tuarob, S., Tucker, C.S., 2013. Fad or Here to Stay: Predicting Product Market Adoption and Longevity Using Large Scale, Social Media Data, ASME 2013 International Design Engineering Technical Conferences (IDETC) and Computers and Information in Engineering Conference (CIE), Portland, OR, USA.

Tucker, C.S., Kim, H.M., 2011. Trend Mining for Predictive Product Design. ASME Journal of Mechanical Design, 133(11), 111008-111008.

Tversky, A., Kahneman, D., 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty, 5(4), 297-323.

Ulrich, K.T., Eppinger, S.D., 2003. Product Design and Development, 3rd ed. McGraw-Hill, New York.

Valente, T.W., 1996. Social Network Thresholds in the Diffusion of Innovations. Social Networks, 18(1), 69-89.

Vicente, L.N., Calamai, P.H., 1994. Bilevel and Multilevel Programming: A Bibliography Review. Journal of Global Optimization, 5(3), 291-306.

Von Hippel, E. 1986. Lead Users: A Source of Novel Product Concepts, Management Science 32 (7): 791–806

Wagner, E.R., Hansen, E.N., 2004. A Method for Identifying and Assessing Key Customer Group Needs. Industrial Marketing Management, 33(7), 643-655.

Wang, X., Camm, J.D., Curry, D.J., 2009. A Branch-and-Price Approach to the Share-of-Choice Product Line Design Problem. Management Science, 55(10), 1718-1728.

Wassenaar, H.J., Chen, W., 2003. An Approach to Decision Based Design with Discrete Choice Analysis for Demand Modeling. ASME Journal of Mechanical Design, 125(3), 490-497.
Wassenaar, H.J., Chen, W., Cheng, J., Sudjianto, A., 2005. Enhancing Discrete Choice Demand Modeling for Decision-Based Design. ASME Journal of Mechanical Design, 127(4), 514-523.
Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications Cambridge University Press, Cambridge, UK.
Watts, D.J., Strogatz, S.H., 1998. Collective Dynamics of 'Small-World' Networks. Nature, 393(6684), 440-442.
Wellman, B., 1997. An Electronic Group Is Virtually a Social Network, in: Kiesler, S. (Ed.), Culture of the Internet. Lawrence Erlbaum, Mahwah, NJ, USA, pp. 179–205.
Wickens, C.D., Hollands, J.G., 1999. Engineering Psychology and Human Performance, 3rd ed. Prentice Hall, NJ, USA.
Wilson, C., 2014. Interview Techniques for UX Practitioners-A User-Centered Design Method. Morgan Kaufmann, Boston, MA, USA.
Wortmann, J.C., Muntslag, D.R., Timmermans, P.J.M., 1997. Customer Driven Manufacturing. Chapman & Hall, London.
Xu, H., Zhou, J., Xu, W., 2011. A Decision-Making Rule for Modeling Travelers' Route Choice Behavior Based on Cumulative Prospect Theory. Transportation Research Part C: Emerging Technologies, 19(2), 218-228.
Yanagisawa, H., Murakami, T., 2007. Emotional Shape Generation System with Exchange of Others' Viewpoints for Externalizing Customers' Latent Sensitivity, ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Las Vegas, Nevada, USA.
Yang, C.C., 2013. An Analytical Methodology for Identifying the Latent Needs of Customers. Total Quality Management & Business Excellence, 24(11-12), 1332-1346.
Zhang, H., Liu, S., 2009. Design of Autonomous Navigation System Based on Affective Cognitive Learning and Decision-Making, International Conference on Robotics and Biomimetics, Guilin, China.
Zhang, Q., Leung, Y.W., 1999. An Orthogonal Genetic Algorithm for Multimedia Multicast Routing. IEEE Transactions on Evolutionary Computation, 3(1), 53-62.
Zhang, Y., Pennacchiotti, M., 2013. Predicting Purchase Behaviors from Social Media, Proceedings of the 22nd International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, Rio de Janeiro, Brazil, pp. 1521-1532.
Zhao, Y., Qin, B., Hu, S., Liu, T., 2010. Generalizing Syntactic Structures for Product Attribute Candidate Extraction, Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, Los Angeles, CA, USA, pp. 377-380.
Zhou, C., Guo, L., 2014. A Note on Influence Maximization in Social Networks from Local to Global and Beyond. Procedia Computer Science, 30(0), 81-87.

Zhou, F., Fang, Z., Xu, J., 2007. Constructing Support Vector Machine Kernels from Orthogonal Polynomials for Face and Speaker Verification, IEEE Fourth International Conference on Image and Graphics, Chengdu, China, pp. 627-632.

Zhou, F., Ji, Y., Jiao, R.J., 2013. Affective and Cognitive Design for Mass Personalization: Status and Prospect. Journal of Intelligent Manufacturing, 24(5), 1047-1069.

Zhou, F., Ji, Y., Jiao, R.J., 2014a. Augmented Affective-Cognition for Usability Study of in-Vehicle System User Interface. ASME Journal of Computing and Information Science in Engineering, 14(2), 021001-021001.

Zhou, F., Ji, Y., Jiao, R.J., 2014b. Prospect-Theoretic Modeling of Customer Affective-Cognitive Decisions under Uncertainty for User Experience Design. IEEE Transactions on Human-Machine Systems, 44(4), 468-483.

Zhou, F., Jiao, R.J., Chen, S., Zhang, D., 2011a. A Case-Driven Ambient Intelligence System for Elderly in-Home Assistance Applications. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 41(2), 179-189.

Zhou, F., Jiao, R.J., 2013a. Hierarchical Bayesian Parameter Estimation for Modeling and Analysis of User Affective Influence, ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference. ASME, Portland, OR, USA.

Zhou, F., Jiao, R.J., Lei, B., 2014c. A Linear Threshold-Hurdle Model for Product Adoption Prediction Incorporating Social Network Effects. Information Sciences, Submitted.

Zhou, F., Jiao, R.J., 2013b. A Nested Multivariate Utility Copulas Approach to Aggregating User Experience Partworths for Aircraft Cabin Interior Design, ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Portland, Oregon, USA.

Zhou, F., Jiao, R.J., 2014. Latent Customer Needs Elicitation by Use Case Analogical Reasoning from Sentiment Analysis of Online Product Reviews. ASME Journal of Mechanical Design, to appear.

Zhou, F., Jiao, R.J., Schaefer, D., Chen, S., 2010. Hybrid Association Mining and Refinement for Affective Mapping in Emotional Design. ASME Journal of Computing and Information Science in Engineering, 10(3), 031010/031011-031019.

Zhou, F., Qu, X., Jiao, R.J., Helander, M.G., 2014d. Emotion Prediction from Physiological Signals: A Comparison Study between Visual and Auditory Elicitors. Interacting with computers, 26(3), 285-302.

Zhou, F., Xu, Q., Helander, M.G., Jiao, R.J., 2011b. Affect Prediction from Physiological Measures Via Visual Stimuli. International Journal of Human-Computer Studies, 69(12), 801-819.

Zhou, F., Xu, Q., Jiao, R.J., 2011c. Fundamentals of Product Ecosystem Design for User Experience. Research in Engineering Design, 22(1), 43-61.

Zhu, X., Ghahramani, Z., 2002. Learning from Labeled and Unlabeled Data with Label Propagation. CMU, CALD Tech Report, CMU-CALD-02-107.

VITA

FENG ZHOU

ZHOU was born in Hangzhou, Zhejiang Province, China. He obtained his Bachelor's degree in Electronic Engineering from Ningbo University, Ningbo, China in 2005; and received his Master's degree in Computer Engineering from Zhejiang University, Hangzhou, China in 2007. He then went to Singapore and received his first Ph.D. degree in Industrial Engineering at Nanyang Technological University in 2012. He is expected to obtain his second doctorate in Mechanical Engineering at Georgia Institute of Technology, Atlanta, GA in 2014. When he is not working on his research, he enjoys watching movies, listening to music, swimming, jogging, and having picnics with friends and family.

His main research interests include engineering design, human factors engineering, human-computer interaction, and information systems. He has published more than 10 journal papers, including IEEE Transactions on Human-Machine Systems, IEEE Transactions of Systems, Man, and Cybernetics, Research in Engineering Design, International Journal of Human-Computer Studies, Interacting with Computers, and ASME Journal of Computing and Information Science in Engineering. In addition, he has published over 20 conferences and 1 book chapter. He has achieved over 460 Google Scholar citations, an h-index of 9, and an i-10 index of 8. His homepage is: https://sites.google.com/site/fzhou35gatech and Google Scholar site is http://scholar.google.com/citations?hl=en&user=u4aZb44AAAAJ.