Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
IDETC/CIE 2015
August 2-5, 2015, Boston, Massachusetts, USA

DETC2015-46764

A MULTIDIMENSIONAL NETWORK APPROACH FOR MODELING CUSTOMER-PRODUCT RELATIONS IN ENGINEERING DESIGN

Mingxian Wang
Research Assistant

Yun Huang
Research Associate

Wei Chen*
Wilson-Cook Professor of Engineering Design

Noshir S. Contractor
Jane S. & William J. White Professor of Behavioral Sciences

Yan Fu
Technical Leader

Department of Mechanical Engineering
Northwestern University
Evanston, IL, USA

Science of Networks in Communities
Northwestern University
Evanston, IL, USA

Research & Advanced Engineering
Ford Motor Company
Dearborn, MI, USA

ABSTRACT

Analytical modeling of customer preferences in product design is inherently difficult as it faces challenges in modeling heterogeneous human behavior and product offerings. In this paper, the customer-product interactions are viewed as a complex socio-technical system and analyzed using social network theory and techniques. We propose a Multidimensional Customer-Product Network (MCPN) framework, where separate networks of "customers" and "products" are simultaneously modeled, and multiple types of relations, such as consideration and purchase, product associations, and customer social networks are considered. We start with the simplest unimodal network configuration where customer cross-shopping behaviors and product similarities are analyzed to inform designers about the implied product competition, market segmentation, and product positions in the market. We then progressively extend the network to a multidimensional structure that integrates customer preference decisions with product feature similarities to enable the modeling of preference heterogeneity, product association and decision dependency. Finally, social influences on new product adoption are analyzed in the same framework by introducing customer-customer relations together with other product-product and customer-product relations. Beyond the traditional network descriptive analysis, we employ the Exponential Random Graph Model (ERGM) as a unified statistical inference framework for analyzing multiple relations in MCPN to support engineering design decisions. Our approach broadens the traditional utility-based logit approaches by considering the dependency among product choices and the "irrationality" of customer behavior induced by social influence. While this paper is focused on presenting the conceptual framework of the proposed methodology, examples on customer vehicle preferences are presented to illustrate the progressive development of the MCPN framework from a simple unimodal configuration to a complex multidimensional structure.

1. INTRODUCTION

Understanding customer preferences, interests, and needs is critically important in developing successful products [1]. Our research is motivated by the need to overcome the limitations of existing quantitative methods for modeling customer preferences in engineering design. Even though utility-based logit models such as Discrete Choice Analysis (DCA)[2, 3] have been widely studied by the design community to guide and optimize design decisions [4-7], there are several major obstacles regarding their use in practical design applications:

- **Dependency of Alternatives.** Standard logit models usually ignore correlations in unobserved factors over product alternatives by assuming observations are independent, i.e., whether a customer chooses one product is not influenced by adding or substituting another product in the choice set, which is often not a realistic situation. Though advanced logit models have been developed to address this issue by introducing certain correlation structures among the error terms, they cannot accommodate dependent decisions explicitly.

* Corresponding author, email: weichen@northwestern.edu

- **Rationality of Customers.** The utility-function based choice modeling approach assumes customers make rational and independent decisions. However, in reality customers influence each other, and their socially influenced decisions can sometimes be considered "irrational." As such, it is reasonable to expect that social effects, such as geographical proximity, communication ties, friendship connections, and social conformity have a heavy influence on customer attitude and behavior.
- **Correlation of Decisions.** Correlated decisions, such as consideration decisions, often involve multiple choices made by the same individual at the same time. It is important to realize that decisions in such situation are often nested within one another. For example, the decision of how many products and what products to consider could be nested. Unfortunately, classical regression models ignore these correlations, and therefore, cannot estimate the influence of the decision outcomes on each other.
- **Collinearity of Attributes.** To evaluate the underlying preference for each product attribute, it is often desirable that preference data has little to no collinearity. However, revealed preference data is very vulnerable to collinearity as the data is drawn from the real market. For example, low price vehicles are more possible to have smaller engine capacity and as a result, low fuel consumptions. However, it is hard to tell whether customers are buying cars because they are low price or because they are fuel efficient. The presence of collinearity implies that the individual contribution of each attribute is difficult to evaluate.

To address these limitations, we propose a *multidimensional network analysis (MNA)* approach, rooted in social network analysis for analyzing complex customer-product relations in support of engineering design decisions. As shown in Fig. 1, using vehicles as an example, customer-product interactions form a complex socio-technical system [8], not only because there are complex relations between the customers (e.g., social interactions) and amongst the products (e.g., market segmentation or product family), but also because there exist multiple types of relations between customers and products (e.g., "consideration" versus "purchase"). Our research premise is that, similar to other complex systems exhibiting dynamic, uncertain, and emerging behaviors, customer-product relations should be viewed as a complex socio-technical system and analyzed using social network theory and techniques. The structure and topological characteristics identified in customer-product networks can reveal emerging patterns of customer-product relations and the interacting effects of product and customer attributes by taking into account the heterogeneities among customers and products.

In literature, network analysis has emerged as a key method for analyzing complex systems in a wide variety of scientific, social, and engineering domains [9]. The approach provides visualization of complex relationships depicted in a network graph, where *nodes* represent individual members and *ties/links* represent relationships between members. Built upon conventional network analysis, social network analysis views social relations in terms of network theory, and the links in the observed network are explained by the underlying social processes such as self-interest, collective action, social exchange, balance, homophily, contagion, and co-evolution [10].

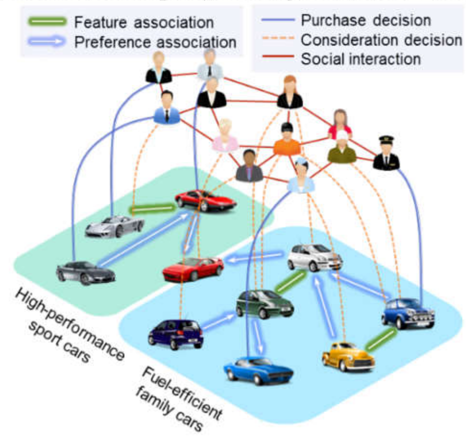


Figure 1：Customer-Product Relations as a Complex Network System

While most existing applications of network analysis are *unimodal* or *unidimensional* that contain a single class of nodes (either human or non-human artifact) and a single type of relation, recent social network research has emphasized on the development of *multidimensional* social networks that include both human and non-human technological elements [11] as nodes and multiple types of relations represented by either non-directed or directed links. Researchers have shown that the network dynamics of a combined human and technology network can shape how people conceive a new technology, as well as whether and how they will use it [12]. In our research, the complex customer-product relations are represented as a multidimensional network (Fig. 1) where products are treated as non-human technological artifacts. Multiple relationships, such as social network relations among customers, association relations among products, as well as preference relations between customers and products, are considered.

Beyond most existing network analyses that are descriptive in nature, our research introduces the *Exponential Random Graph Model (ERGM)* as a unified statistical inference framework for MNA. ERGM is increasingly recognized as one of the central approaches in analyzing social networks [13-15]. ERGMs account for the presence (and absence) of network links and thus provide a model for analyzing and predicting network structures. ERGMs have several advantages: (1) Network links are modeled to be interdependent in ERGM rather than assumed to be independent, (2) ERGMs can incorporate binary, categorical, and continuous node attributes to determine whether they are associated with the formation of network links, (3) ERGMs are capable of characterizing local and global network features; (4) ERGMs can be applied in flexible ways to many different types of network and relational data; (5) Data used for fitting ERGMs can be cross-sectional or longitudinal (change with time); and a dynamic model can be built to study the emergence and dynamics of a network.

This paper employs MNA for the study of customer-product relations as a complex Multidimensional Customer-Product Network (MCPN) in the context of engineering design. While the proposed MCPN is widely applicable for analyzing and predicting any type(s) of preference relations between customers and products, the detailed methodological development in this paper is limited to modeling customer consideration preferences among a set of competing products and examining how the results can inform the values of product attributes in design. The rest of the paper is organized as follows. Sec. 2 introduces the technical background and recent accomplishments in social network research. Sec. 3 describes the development process of MCPN progressively from a unimodal structure to a multidimensional structure with multiple types of nodes and links. Sec. 4 develops two network implementations using the vehicle preference data in China market to illustrate the network approaches - a descriptive approach on the unimodal structure and a statistical inferential technique using MCPN. We have chosen these two implementations because each enables us to gain new insights into the issues that we have not addressed specifically in the traditional Discrete Choice Analysis. Finally, Sec. 5 discusses the pros and cons of the MNA approach and the opportunities for future research.

2. TECHNICAL BACKGROUND

2.1 Network analysis in product design and market study

Network analysis has received considerable interest in product design and market study. In product design, network analysis has been used to characterize a complex product as a network of components that share technical interfaces or connections. Using the network metrics such as "centrality", Sosa et al. [16] defined three measures of modularity as a way to improve the understanding of product architecture. Based on Sosa's work, Fan et al. [17] developed a bottom-up strategy for modular product platform planning. A recent work by Sosa [18] found that proactively managing the use of network structure (such as hubs) may help improve the quality of complex product designs. Network analysis has also been applied to studying designers' network for understanding organizational behavior [11] and improving multidisciplinary design efficiency [19]. In market study, text-mining apparatus has been integrated into a network analysis framework to understand customers' top-of-mind associative network of products based on the large-scale, customer generated data posted on the Web [20]. However, the constructed product-feature network is unidimensional, without including customers and their relations to products in the same network. In contrast to the existing unimodal product network analysis approaches, our multilevel multidimensional customer-product network (MCPN) is built with both product and customer nodes, together with product feature associations and customer social network, to understand how customer decision-making interacts with product attributes and how social influence affects individual decisions for new products.

2.2. Modeling the impact of social influence

Modeling the impact of social influence has received increasing attention in product design [21]. A comprehensive study of how peer influence affects product attribute preference was provided by Narayan et al. [22] who modeled three different mechanisms of social influence. By combining traditional conjoint analysis on product features with peer influence, their work showed that peer influence causes people to change perspective on product importance, and that some product attributes are more sensitive to change than others. However, the approach requires a strict format of survey data to evaluate the attitude change before and after exposure to peer influence.

In modeling social influence in customer vehicle choices, a simulation-based approach has been developed in our earlier research to capture the dynamic influence from social networks on the adoption of hybrid electric vehicles [23]. The social network impact is captured via introducing "social influence attributes" into the discrete choice utility function. The effects of these attributes are assessed through the social network simulation, where the network was constructed based on the "social distances" measured by the dissimilarities of customers' social profiles. Similar assumptions of social influence spreading over a small world network have also been found in [24, 25]. In this research, a multidimensional network approach is proposed to measure simultaneously customer-customer social interactions together with customer-product preference relations for assessing social impact on preference decisions. A simulation-based social network construction approach, similar to [23], is applied to convert customer attribute vectors into relational data that takes into account interdependence of attributes and the interactions between customers and products.

2.3. Advances in Social Network Analysis

In the past decade, social network scholarship has made a concerted effort to move from describing a network to developing techniques that explain the emergence and dynamics of networks. Development of analytic techniques to explain the emergence of networks is often motivated by *multitheoretical multilevel (MTML)* models [10]. Social network models are multi-theoretical because of a growing recognition among social networks researchers that the emergence of a network can rarely be adequately explained by a single theory. Therefore, these models combine disparate theoretical generative mechanisms, such as self-interest, collective action, social exchange, balance, homophily, proximity, contagion, and co-evolution. Multilevel network data categorizes nodes into different levels, and the network links represent relationships between nodes within and across different levels. A *unimodal* (one-mode) network can be defined within each level, and a *bipartite* (two-mode) network can be defined between nodes from two levels. These models are also multilevel because the emergence of a network can be influenced, for instance, by theories of self-interest that refer to characteristics of actors (at the individual level), theories of social exchange that describe links between pairs of actors (at the dyadic level), theories of balance that explain the configuration of links among three actors (at the triadic level), and theories of collective action that explain configurations among larger aggregates of actors (at the group or network level). The network configurations at multiple levels capture the interdependency among links and between links and nodal attributes to represent the underlying behavior patterns.

ERGMs provide the statistical inference framework for MNA. Technically, we can define matrix **Y** as a random graph in which rows and columns represent customers and products, respectively. $Y_{ij} = 1$ refers to a relation, such as the preference decision between customer i and product j, and 0 otherwise. ERGMs have the following form:

$$P_{\theta}\left(\mathbf{Y}=\mathbf{y}\right)=\frac{1}{c\left(\boldsymbol{\theta}\right)}\exp\left\{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{z}\left(\mathbf{y}\right)\right\}, \qquad (1)$$

where (i) **y** is the observed network, a random realization of **Y**; (ii) **z(y)** is a vector of network statistics corresponding to certain network configurations in **y**, and the settings of product and consumer attributes; (iii) **θ** is a parameter vector indicating the effects of the network statistics; (vi) c is the normalizing constant that ensures the equation is a proper probability distribution. Eqn. (1) suggests that the probability of observing any particular graph (e.g. MCPN) is proportional to the exponent of a weighted combination of network characteristics: one statistic is more likely to occur if the corresponding is positive. Our research aims to interpret the meaning of these parameters (network effects) in order to understand customer-product relations for product design. More details are provided in Sec. 3.3.

3. A MULTIDIMENSIONAL NETWORK APPRAOCH FOR PREFERENCE MODELING

3.1. The Multidimensional Customer-Product Network (MCPN) Framework

In this paper, we recast the problem of modeling customer preferences as network modeling of customer-product relations. We view engineering products as an inherent part of the expanded social network along with human actors. Fig. 2 describes the structure of the MCPN framework, which is characterized by two classes of nodes at two layers ("product" and "customer") and multiple types of relations within and between the two layers.

The ***product layer*** contains a collection of engineering products **P** (e.g., vehicles, electronics and appliances, software). Products are connected by various links which can be either directed or non-directed. *Directed links* often involve product hierarchy or preference, while *non-directed links* imply product similarity or association. Product attributes or features, quantitative (e.g. fuel efficiency, horsepower) or qualitative (e.g. safety, styling), can be taken into account as nodal attributes. Similar attributes/features between products are represented as association links in the product network. Alternatively, product associations can be identified by their co-consideration relations from customers. The ***customer layer*** describes a social network consisting of a customer population **C** who make decisions or take actions. Each customer has a unique profile (e.g., socioeconomic attributes, purchase history, etc.) which potentially affects customer preference decisions. Links between two customers represent their social relations, such as friendship or communication. The structural tendencies of these social relations reflect the underlying social processes for creating and maintaining links such as homophily and proximity [10]. ***Customer-product relations*** are indicated by various human activities such as purchase and consideration decisions. The *customer-product links* are created between two sets of nodes from two adjacent layers, representing customer preference. As shown in Figure 2, if a customer purchase a product, there will a solid link between the customer and product nodes. If customer considers a product, the link between the two nodes are marked as a dashed line. As noted, a customer can consider several products at the same time while the final purchase is only one or none. These preference links can be flexibly constructed by various sources of data, e.g., survey data, transaction data, and user-generated text data.

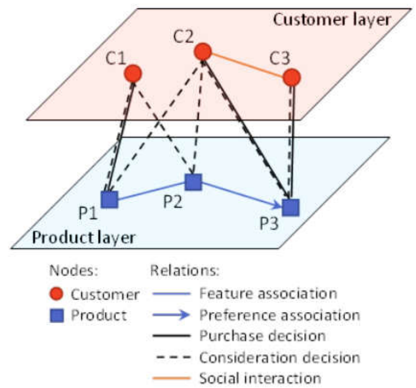


Figure 2: Multidimensional Customer-Product Network

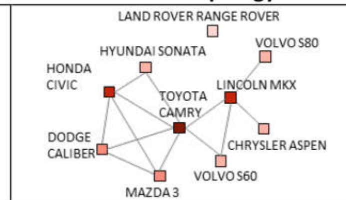
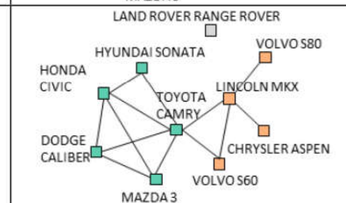
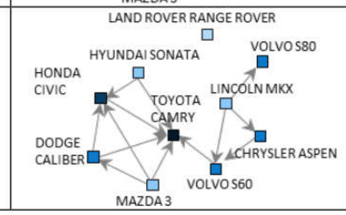
As seen, the proposed MCPN framework can capture rich information on dependency in a complex socio-technical system so as to assist product design decision-making. A combined analysis of all relations mentioned above allows designers to evaluate product decisions not in isolation, but with expectation that the market system will react to the planned decisions, and any design change may easily affect other connected entities across the network in ways that were initially unintended.

3.2. Unimodal Network Analysis of Product Associations

Our development of MCPN started with the *unimodal network analysis* to a single layer network with only product nodes and associations. The unimodal network can be viewed as a compressed but simplified version of the more complicated bipartite (customer-product) networks by projecting it to a single layer [9]. The unimodal network enables designers to explore the use of descriptive metrics in identifying *aggregated* product associations that can reveal the implied product similarity and diversity, product market competence, product market segmentation, and other opportunities for design improvements.

The links in a product association network can be constructed in many ways. For example, using the customer preference data, a *customer-driven product association network* can be established, where the links between products reflect the proximity or similarity of two products in customers' perceptual space. Alternatively, a *feature-driven product association network* can be established with the help of product specification data, where the association between products can be determined by measuring the similarity of product attributes/features from designers' point of view.

Table 1: Examples of Network Structural Analysis for Analyzing Customer-Driven Product Associations

	Network Analysis	Solution Techniques	Network Topology
Centrality	**Centrality** involves the identification of the 'most competitive' products in the network [9]. We assume that more central (or more connected) products have higher levels of survivability in market competitions as a result of its structural advantageous.	Measuring centrality can be based on various properties of a node, e.g., number of direct connections to all other nodes (degree), minimum distance to all other nodes (closeness), and maximum occurrence on the path of two other nodes (betweenness) [9].	
Community	**Community** refers to the occurrence of groups of nodes that are more densely connected internally than with the rest of the network [26]. If appropriate communities are detected, the network can be collapsed into a simpler representation without losing much useful information.	The modularity maximization method [26] can be used as the objective function to capture the quality of a network structure. The problem is solved as an NP-hard optimization problem.	
Hierarchy	**Hierarchy** is formally defined as a strict partially ordered set which can be represented as a directed network [27], where each element of the set is a node and the partial ordering (P1<P2) gives an edge from P1 to P2. The directed link reflects customers' aggregated preference across the population.	To find local hierarchies of nodes, centrality metrics can be applied as well to a directed network configuration. To bring global order to the nodes, heuristic search algorithms, e.g. Google's PageRank [28], can be employed to find the best hierarchy in a polynomial time.	

The descriptive network analysis involves the computation of topological measures to assess the position of nodes and the implication of structural advantages. Examples for analyzing customer-driven product associations are provided in Table 1. *Centrality* [29] measures a product's competitiveness, indicated by its level of connectivity to other products. *Community* [26] analysis identifies products with close connections as a market segment. Network *hierarchy* [30] is illustrated by the directed links which encode preference rankings.

Although the unimodal network approach can describe interdependencies in relational data, the method cannot provide quantitative assessment of product attributes for a particular group of customers. Further, the unimodal network analysis studies customers' averaged (aggregated) preference across the population. Advanced network modeling approaches that capture disaggregated preference behaviors of individual customers are needed as examined next.

3.3. Analyzing Multidimensional Network Considering Product Associations

To model heterogeneous customer preferences in products with close associations, we integrate the product association links with customer-product preference relations as a multidimensional network (see Fig. 3). By introducing the information from the second mode (i.e. customers), we aim to develop a network model capable of capturing customer preference heterogeneity and multiple dependent decisions considering product feature associations. The information obtained can be used to identify the right product configurations for a targeted group of customers in product design.

Beyond existing network approaches that are mostly descriptive in nature, we use ERGM as a unified statistical framework to analyze the MCPN. In ERGMs, the observed network is considered one realization of an underlying probabilistic distribution, without assuming the independence of nodes or links. A local topological configuration in the network, i.e. a set of connected nodes and links, is regarded as an exploratory variable representing the structural features of potential interest. Networks in the distribution are assumed to be "built up" from the localized patterns represented by the structural features. ERGM literature has established more than 20 different types of effects [13] for describing the various forms of dependence that exist in the relational data within social networks. Examples of effects, their configurations, and interpretations are provided in Table 2.

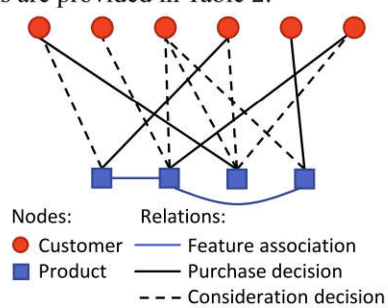
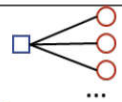
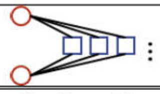
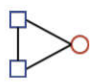


Figure 3 : Multidimensional Network Considering Product Associations

The network effects fall into three categories: *pure structural effects* are related to well-known structural regularities in the network literature; *attribute-relation effects* assume the attributes of products/customers can also influence the formation of network links in addition to the structural endogeneity in the network; the product association relations can be characterized by the *cross-level effects* that integrate customer preferences with product similarities.

Table 2: Examples of Interpretations of Network Effects in MCPN

Pure Structural Effects	Configuration	Interpretation
[A] Density		This effect captures the baseline propensity of forming a link. It is similar to the intercept in a regression model
[B] Alternating k-stars for products		This effect measures the dispersion of the degree distribution. Alternatively, it can be thought as a test of the "rich get richer effect". **Example:** A positive parameter indicates that the network links are centralized around a few high-degree nodes of products.
[C] Alternating k-cycles for customers		This effect captures the propensity of customers to engage in closed structures. **Example**: Two customers considered the same product also consider some other products together
Attribute-Relation Effects	**Configuration**	**Interpretation**
[D] Main effect		This effect captures whether the binary attribute or higher scores on a continuous attribute tend to express more links. **Example**: A significant negative parameter for vehicle fuel consumption means fuel efficient cars are more likely to be considered by customers.
[E] Interaction effect		This effect captures the interaction of the nodes between different types. **Example**: A significant positive coefficient for family size of customers and vehicle size of products suggests customers from large families tend to consider large size cars.
Cross-level Effects	**Configuration**	**Interpretation**
[F] Association based closure effect		This effect captures whether a closed structure is more likely to occur involving two product nodes with an association link. **Example**: A negative significant coefficient means that customers do not tend to consider two cars with many common features at the same time.
■Product w/ attributes ● Customer w/ attributes □ Product w/o attributes ○ Customer w/o attributes		

Once the network effects of interest are identified by designers, their significance can be determined by estimating the model parameters of an ERGM via likelihood maximization, given the observed network data. As the exact maximization of the likelihood function requires a summation over all possible configurations of the network and is computationally demanding, approximation techniques (e.g., maximum pseudolikelihood [31], Markov Chain Monte Carlo maximum likelihood [32]) can be employed to find the estimates of effects.

Compared to a unimodal network (Sec. 3.2), a multidimensional network is a more natural way to model relations between two different classes of nodes and non-hierarchical association relations between products. Moreover, its capacity to preserve two types of nodes allows researchers to parse out the unique contribution of different types of nodes to the overall network structure. Its ability to integrate product networks and customer-product relations allows researchers to model interdependent product relations and correlated preference decisions explicitly, without specifying complicated error structures as often done in DCA.

3.4. Analyzing Multidimensional Network incorporating Social Influence

To account for the effect of social influence on customer preference decisions, we further expand the multidimensional network structure to simultaneously measure within-layer social relations, within-layer product associations, and between-layer customer-product relations (Fig. 4).

The proposed multidimensional network allows the evaluation of both the "peer effect" and the more general "crowd effect" [33], depending on how product associations and social relations are defined. Relations between customers are used to model "peer effect" on customer attitudes and preferences. The term "peer" has a broad meaning which may include "friends," "neighbors," "experts," "relatives" or even "online reviewers" with whom customers may exchange information about new products. The preference hierarchies among products, as defined in Sec. 3.2, can be used to capture the effect of "social crowd". The evaluation of effects is done by assessing the structural tendencies of networks informed by social influence theories (Table 3). Using ERGMs, one can quantify the effects of social influence by statistically estimating the extent to which structural tendencies implied by social theories influence the probabilities of the observing network. Similar to the network effects in Table 2, customer and product attributes can be incorporated into the social influence structures for investigating how social influence varies across customers and products.

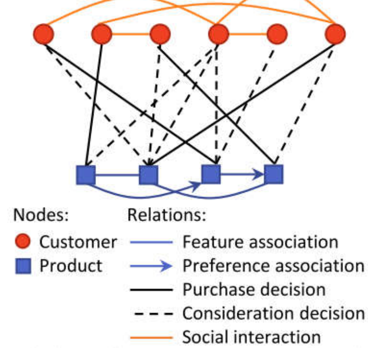
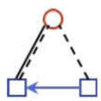
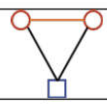


Figure 4: Multidimensional Network Consideration Social Interactions

Due to the complexity of data collection, customer social network data is often not collected in consumer surveys. An alternative is to construct social relations through network simulations [23], based on certain hypotheses of network structure and "social distance" measured by the collected customer profiles. For example, based on the theory of homophily [34], we can assume that two nodes with shorter social distance (similar customer attributes) are more likely to be connected. Unlike the prior research that incorporates social influence as customer attributes, this research employs the ERGM to assess the social influence effects. In theory, one should draw more reliable conclusions based on the results from the network approach, because of its capability of handling correlated node attributes and interdependent link relations, and thus avoids faulty inferences on covariates [35].

Table 3: Examples of Social Influence Effects in Multidimensional Network

Social Influence Eff.	Config.	Interpretation
Crowd effect on purchase		When comparing two products under consideration, a customer is more likely to purchase the one favored by the majority of customers.
Peer effect on purchase		Customers tend to purchase the product that their "peers" recommended, either through use or discussion.
□ Product w/o attributes		○ Customer w/o attributes

4. CASE STUDY – VEHICLE PREFERENCE MODELING

4.1. Using Unimodal Network for Modeling Vehicle Associations and Hierarchies

Two implementations on modeling customer vehicle preference in the growing China market are presented to demonstrate the proposed methodology. In the first implementation, we demonstrate the unimodal network analysis (Sec. 3.2) for identifying aggregated product associations and hierarchical preference relations. Beyond existing literature, our work utilizes both consideration and purchase data in market surveys to derive relationships among vehicle products for understanding customer preferences and product competitions. We develop two types of product association networks – a *vehicle association network* with *unidirected* links showing the similarity of products, and a *vehicle hierarchal network* with *directed* links indicating preference hierarchies.

The two vehicle association networks are constructed using 2013 New Car Buyers Survey (NCBS) data provided by an independent research institute in China. The dataset contains 49,921 new car buyers who considered and purchased from a pool of 389 vehicle models in 2013. Both the set of considered vehicles and the final purchase are recorded for each customer. Customer demographics and product information are also reported by respondents.

The vehicle association network is created to aid the analysis of customer consideration decisions by linking any pair of vehicles if both vehicles are considered by the same consumer in his (her) consideration set. The association link is viewed as a form of similarity or closeness between any two vehicles in customers' minds. The link strength is quantified by *lift* to reflect how often the two products are compared by a population of customers. The lift between product i and product j is defined as the probability of co-consideration over the probability that they are being considered individually. The probability value is approximated by the percentage of product (co)occurrence recorded in NCBS data.

$$lift(i,j) = \frac{\Pr\{\text{co-consider } i \text{ and } j\}}{\Pr\{\text{consider } i\}\cdot\Pr\{\text{consider } j\}}. \qquad (2)$$

To prune the network links, a thinning threshold at 1 is chosen for the lift value, because a lift greater than 1 has a precise statistical meaning showing a positive association between the two products [36]. For example, Honda Guangzhou Odyssey and Mazda FAW 8 are positively associated, as shown in Fig. 5(i). The association link implies that the two products have a high chance of being co-considered. From the customer's perspective, it means that a customer considers Odyssey is also very likely to consider Mazda 8 at the same time.

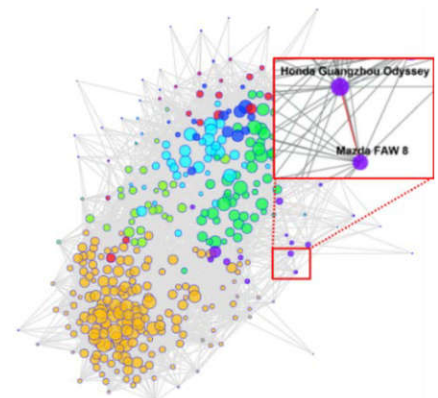


Figure 5(i). Centrality and Community in Vehicle Association Network based on NCBS 2013. Nodes are Sized by Network Degrees and Colored by Network Communities

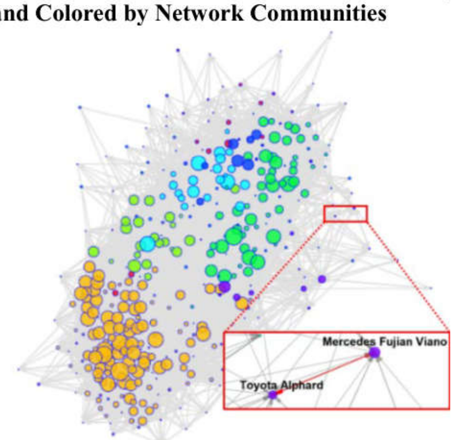


Figure 5(ii). In-degree Hierarchy in Hierarchical Preference Network based on NCBS 2013. Nodes are Sized by Network In-degrees and Colored by Network Communities.

As a measure of network *centrality*, the node degree calculates the number of links attached to a node. In the vehicle association network, products with a higher degree centrality are those frequently co-considered with many other vehicles by customers. Examples of high-degree centrality vehicles include GM SGM Chevrolet Sail, Audi FAW Q5, and Kia Dongfeng Yueda K2. One interesting observation is that most of the high-centrality vehicles are also among the most popular vehicles considered by customers, though the two quantities are not equivalent in definition. Another observation is that the node degrees is not uniformly distributed such that some vehicles are considered more frequently than others.

For the constructed vehicle network, the product *community* analysis is employed following Newman's modularity method to determine groups of interconnected vehicles. In Fig. 5(i), the seven identified communities are marked in different colors. The product communities inform designer the marketing coverage of a brand family and marketing competence across several brands. For example, the yellow community includes most domestic entry-level sedans (e.g., BYD F6, Chery QQ, etc.), while the green community is featured by premium SUVs by foreign manufacturers (e.g., Jeep Grand Cherokee, Land Rover Discovery, etc.). It is also observed that a product line's marketing success is highly influenced by its product positioning strategy. The successful product lines in the market generally cover more network communities. For example, as two marketing leaders in China, Volkswagen and GM have covered 6 out of the 7 network communities, implying a great diversity of their vehicle products across multiple segments.

As a refinement to the above undirected network, a *directed network* is constructed where a link direction is determined through both consideration and purchase data in NCBS. If for any pair of vehicles, a customer considers both vehicles but chooses one over the other, the link direction will point towards the purchased vehicle. The lift metric shown in Eqn. (3) is slightly modified to accommodate the evaluation of directed link strength.

$$lift\left(i \to j\right) = \frac{\Pr\left\{\text{co-consider } i \,\&\, j, \text{purchase } j\right\}}{\Pr\left\{\text{consider } i\right\}\Pr\left\{\text{purchase } j\right\}} \quad . \quad (3)$$

Again, the links are trimmed to highlight positive associations. The link direction captures the preference hierarchy between the two linked products. For example, a bi-directional (mutual) link between Toyota Alphard and Mercedes Fujian Viano can be interpreted as the intense competition between the two products (Fig. 5(ii)), because either vehicle model attracts a great percentage of customers in consideration. Nevertheless, Viano gains a slight upper hand in market competition, because the strength of the link in that direction is stronger.

With a directed network, graph metrics indicating node *hierarchy*, such as node in-degree, can be computed to reflect customers' aggregated preferences across the population. The *in-degree* of a node computes the number of incoming links pointed to that node. A node with a high in-degree value implies the corresponding vehicle is very likely to be considered with other vehicles and is also more preferred in customer choice (purchase) decisions. For example, Audi FAW Q5 and Ford Kuga are popular vehicles in choice, which are ranked high in both degree centrality and in-degree hierarchy. In contrast, Volvo V40 and Ford Edge have been frequently considered (high degree centrality in undirected network), but fall behind in customers' final choices (low in-degree hierarchy).

Our illustrative example shows that descriptive network analysis may serve as a useful tool to determine product positioning and product priorities in the phase of design planning. C*entrality*, *community*, and *hierarchy* allow designers to uncover the root-causes of the differences in vehicle sales under a specific market. These efforts may reveal issues that a design team could work on, e.g., product recognition (low centrality rank), coverage and diversity of product lines (products not appearing in certain communities), product competence (several vehicles in the same community), and product configuration (low hierarchy rank), etc.

While analyzing the structural information of a unimodal network can be useful in describing product associations, there is a need for an approach to quantitatively evaluate customer heterogeneous preferences while addressing issues such as dependent alternatives, multiple decisions, social influence, and correlated observations. To demonstrate such capabilities of a network model, our next example employs ERGM in the MCPN framework with various nodes, relations and attributes included.

4.2. Using MCPN for Modeling Luxury Vehicle Preferences in Central China

Our second implementation demonstrates the use of inferential network technique (ERGM) for analyzing the vehicle MCPN framework (Secs 3.3 & 3.4). This network implementation also draws from the 2013 NCBS data to understand customer preference trends in China. With a focus on the luxury vehicle market, we examine respondents who live in the central provinces of China and consider only luxury imported vehicle models in their decision journey. This focused interest results in a subset data of 378 customers and 65 luxury vehicle models for modeling and evaluation. As reported by McKinsey, the top reasons for Chinese customers to choose a luxury vehicle are: "reflection of social status", "self-indulgence" and "business credibility". Therefore, we expect that socially influenced decisions are more common in luxury vehicle buyers in China. In addition, the Chinese auto market is renowned for its complexity and volatility. Strong regional differences exist as a result for brand accessibility and lifestyle needs. Because of these hidden reasons beyond the functionality and design of a vehicle itself, quantifying the attractiveness of a vehicle attribute in such conditions becomes even more difficult.

The proposed MCPN integrates a feature-driven product association network, a customer-product network, and a customer social network as a unified entity for analysis. The implementation of the proposed approach goes beyond the descriptive analysis and consists of three major steps: network construction, ERGM specification, and ERGM interpretation; each of these steps is explained in the remaining of this section.

4.2.1. Data Transformation & Network Construction

1) Product Associations. Depending on the product complexity and the purpose of analysis, product associations can be built using either the "complete set of features" or "subsets". In this example, association links are constructed using three vehicle categorical attributes -- body type, brand, and size segment. The product association link is viewed as a form of similarity in the "overall styling" between any two vehicles; whether they are comparable in terms of price level, safety or performance, the association link is indicative of how similar the two vehicles visually "look like". By converting body type, brand, and size attributes as product associations, our emphasis in network inferential analysis using EGRM is on the influence of other vehicle attributes on customer preferences. Within the association network construct, the Gower's coefficient [37] is calculated to determine the existence of a link between any product pair. Gower's coefficient has the capability to appropriately handle continuous, ordinal, nominal and binary variables as inputs. After several experiments, a global thinning threshold is then decided to produce reasonably dense network to ensure reliable ERGM estimates and avoid network degeneracy in model estimation [13].

2) Preference Relations. We use the bipartite links between a product and a customer to model customers' consideration decisions over vehicle models. The structure of these links is precisely defined by NCBS data. In the survey, respondents are asked to report a list of vehicles that they seriously considered, including the purchased one. The number of consideration number ranges from 1 to 3. No customer listed more than 3 vehicles, even though the actual number might be higher.

3) Social Relations. Unlike the product association links which can be flexibly determined, the social links between customers have more specific meanings in social theories. Continuing our previous work on network simulation [23], a social space is constructed based on customer geographical locations and selected social attributes (age, income, education). Based on the homophily assumption that two customers with shorter distance in the social space are more likely to be connected, a global threshold is chosen to determine if a social link exists or not. To mimic the properties of real world networks, we then adjust the social links using the small-world model [24, 25] to assure the high transitivity ("one's friends are likely to be friends") and low average path length ("six degrees of separation" between any two individuals). The small world mechanism provides a viable way to represent social links through both close and distant connections, implying that customers are not only influenced by their nearest neighbors in their social space but also a small number of remote contacts outside their regular social proximity.

Integrating the three types of network relations together, a visualization of the construction process for the MCPN structure is presented in Fig. 6. The complexity of network progressively increases from product association only in Fig. 6(i), to adding customer-product relations in Fig. 6(ii), to adding the customer-customer relations in Fig. 6(iii). As noted, we only include one type of preference link (consideration) and one type of product association link (feature-driven) for demonstration. All links are binary-valued and undirected.

4.2.2. Specification of ERGMs for Multidimensional Networks

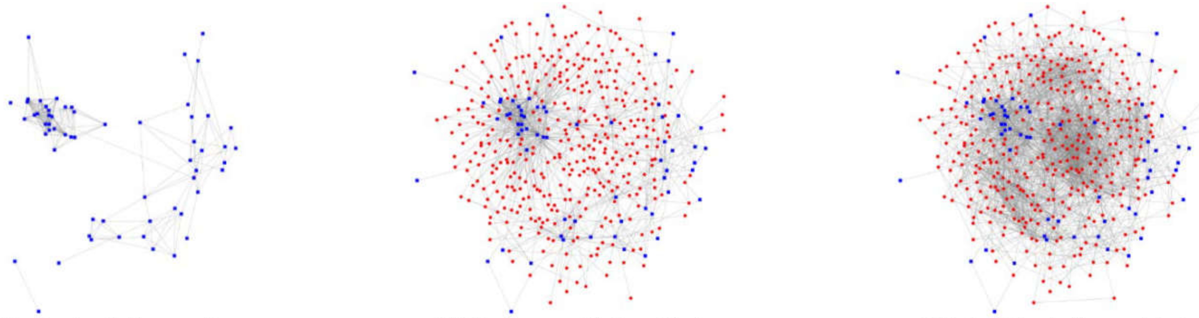
With the constructed MCPN structure, the conditional form of ERGM is employed to address the question of how one or more dimensions of networks would affect the structures of the other networks. Specifically, our research question is about how product associations and social relations may impact customers' consideration decisions. As presented in Table 4, the examined network effects are restricted to a subset of cross-level configurations and product/customer attributes of different forms. The choice of which network effect to include depends on the social theory, hypothesis, and the specific research questions to answer. Nevertheless, the demonstrated example serves as guidance for possible effects to consider in vehicle preference modeling for engineering design.

4.2.3. Comparisons and Interpretations of ERGMs

Estimating the model coefficients for ERGM network effects is equivalent to fitting a model that gives maximal support to the data. However, the maximum likelihood estimates cannot be derived analytically due to the intractable constant in Eqn. (1) for a reasonable number of nodes. Thus, we employ a stochastic approximation [38] that relies on MCMC simulations of graphs.

We compare three model specifications based on the same data set to highlight the benefits of the ERGM approach. Model 1 formulates a bipartite ERGM analogous to a logistic model that contains only the dyadic actor-relation effects (only customer-product relation, but no product-product or customer-customer relation). This model allows the testing of influencing customer/product attributes in customer preference decisions, assuming that endogenous structural processes do not exist. Model 2 parameterizes a multilevel ERGM similar to Model 1 but with the addition of the pure structural effects and the cross-level product association effect. By comparing Model 2 and Model 1, we can test whether the addition of these structural effects may modify some of the attribute-relation effects in explaining customer preferences. The specification of Model 3 is the most complete model that includes all three types of ERGM effects. With the integration of the cross-level social influence effect, peer influences on preference decisions can be evaluated together with other product attributes, customer demographics, and structural patterns within the same model.

The interpretation of Model 1 is similar to that for a logistic model. The significant positive *turbocharger* and *engine capacity* indicate that the presence of the turbocharger and the increased size of the engine would increase the probability for a customer to consider a particular vehicle model. The statistically negative *first-time buyer* suggests that first-time buyers are unlikely to enter the luxury vehicle market even though three out of four new cars are purchased by first-time buyers in China. The *fuel consumption* has a significant positive coefficient, meaning that fuel economy is less important for customers who decide to purchase a luxury vehicle. The insignificant effect of vehicle *price* implies that price is not the main factor to characterize a popular product model from other models in the luxury vehicle market. Likewise, the decision of how many vehicles to consider is less relevant to the household income, as seen by the insignificant *income* in the table. As noted, most model coefficients in Model 1 agree with our prior understanding about China's luxury market. This means that including attribute-related effects alone can capture an essential component of the process underlying the MCPN structure.



(i) Product relations only **(ii) Preference links added** **(iii) Social relations added**

Figure 6: Construction of Multidimensional Network using NCBS Data. Products as Blue Squares and Customers as Red Disks.

In Model 2, the addition of the pure structural effects and the cross-level association effect considerably changes the interpretation of the underlying preference data. The significant positive *alternating k-stars for products* indicates the dispersed distribution of product nodes such that customers' decisions mostly concentrate on only a few vehicle models in the market. In contrast, the degree distribution is more centered for customer nodes, as shown by the negative *alternating k-stars for customers* coefficient, because customers only consider a limited number of vehicles (1-3) in NCBS data. The *association based closure* effect is an indicator of how likely a customer may consider two vehicles which are visually similar. The significant positive coefficient means most people would judge a vehicle by its appearance and consider multiple vehicles with "similar look". Concerning the attribute-relation effects, all the product effects (*turbocharger*, *engine capacity*, *fuel consumption*) generate smaller coefficients in magnitudes to their counterparts in Model 1 and the *engine capacity* is no longer significant. The customer effects of *first-time buyer* and *income* become more obvious, partly because the number of decisions (degree of customer nodes) has been controlled by the *alternating k-starts*.

The coefficients of Model 3 are largely consistent with those in Model 2, except that the previously insignificant *income* becomes significant. The significant positive *peer influence* indicates that a customer is likely to become "irrational" in decision making and simply considers what his/her peer has considered. Modeling the peer influence is a unique contribution of our work as such effect cannot be modeled both theoretically and computationally without the MCPN framework.

By comparing the above three models, several interesting findings can be summarized about the preference modeling in a multilevel network context. First, including the attribute-relation effects alone (Model 1) can explain a large part of the formation of preference links. This observation is consistent to the foundational theory of many attribute-based preference modeling approach, such as DCA. Second, a model with only attribute-relation effects but no other relevant structural effects may ignore some of the underlying social structures represented by the structural patterns; therefore, such a model may produce biased results even if a researcher is only interested in a subset of product/customer attributes. Finally, the peer influence effect (Model 3) introduces another layer of dependencies between two customers into the structure of the network. The significant positive estimate reflects the importance of social influence in explaining customer behavior and modeling product demand. Overall, the results of this example suggest that the nodal attributes (customer and product attributes) and network structures (product associations, social influences, and other underlying effects) are both indispensable elements and play together in shaping the decision behaviors of customers.

Table 4: Comparison of three specifications of ERGMs.

	Model 1		Model 2		Model 3	
Effects	Est. Coeff	(Std. Err)	Est. Coeff	(Std. Err)	Est. Coeff	(Std. Err)
Pure Structure Effect						
Density	**-7.0138**	**(0.399)**	**-3.0057**	**(0.659)**	**-2.8950**	**(0.656)**
Alternating k-stars for products			**0.8549**	**(0.221)**	**0.8818**	**(0.222)**
Alternating k-stars for customers			**-4.5553**	**(0.463)**	**-4.5640**	**(0.446)**
Attribute-Relation Main Effect						
Price paid to the dealer (in 100K RMB)	-0.0333	(0.019)	-0.0139	(0.019)	-0.0103	(0.018)
Turbocharger dummy	**1.2776**	**(0.110)**	**0.9124**	**(0.129)**	**0.7976**	**(0.118)**
Engine capacity (in cc)	**0.2724**	**(0.135)**	0.0767	(0.127)	0.0378	(0.123)
Fuel consumption (in L/100km)	**0.1578**	**(0.039)**	**0.1147**	**(0.036)**	**0.0993**	**(0.032)**
First-time buyer dummy	**-0.2274**	**(0.096)**	**-1.0463**	**(0.222)**	**-1.0295**	**(0.224)**
Monthly household income (in 100K RMB)	0.0027	(0.002)	0.7085	(0.373)	**1.0002**	**(0.386)**
Cross – Level Effect						
Association based closure			**0.9287**	**(0.156)**	**0.9397**	**(0.127)**
Peer influence					**0.2746**	**(0.016)**
Bolded coefficients are different from null at the 95% confidence interval.						

5. DISCUSSION AND CONCLUSION

While Discrete Choice Analysis (DCA) has been widely used to predict the influence of design decision on customer preference and firm profit, in this paper, we introduce a new network approach that enhances traditional DCA models for analyzing customer-product relations in supporting engineering design decisions. We demonstrated the progression of a simple unimodal network that contains only product associations, to a multidimensional network that considers product associations together with customer preference decisions, and finally to a more complete multidimensional structure that integrates product associations, customer social influence, and preference decisions as one network entity.

The descriptive network analysis as presented in the unimodal network example offers a convenient tool to summarize key facts about the customer preference data. Through descriptive network measures, nodes can be clustered into subsets (community) or organized in ranks (centrality, hierarchy) to reflect structural positions in a network. When complex product association relationships are converted into market segments and competitive rankings, designers can better monitor product positions within a brand or between brand competitors. The inferential network analysis with ERGM as illustrated in MCPN configuration enables the detailed modeling of both the network structures and customer/product attributes in a rigorous statistical sense. Conceptually, the advantages of the ERGM for multidimensional network analysis (MNA) over traditional logit models for modeling customer preferences can be summarized as follows:

- Product associations can be modeled explicitly. In ERGM, product alternatives are no longer mutually exclusive, but interdependent in a network structure to influence customer's preference decisions.
- Evaluation of social influence is enabled. By constructing customer social links in the customer layer, ERGM allows the social network effect to be statistically assessed and compared with other factors within a single model.
- Nested decisions can be analyzed through structural modeling. The model estimates can uncover not only a customer's taste for a particular product, but also the relationship between several preference decisions as well as the number of decisions made, as represented by the correlated structural effects.
- Correlated product/customer attributes are possible. Since ERGM assumes the observed network as a single realization from a multivariate distribution, no I.I.D assumptions are necessary over the explanatory variables. Correlated product/customer attributes can be entered as structural terms and evaluated simultaneously.
- Coefficient estimates are highly interpretable and the ERGM results can be easily integrated into an engineering design optimization problem. The model estimates in ERGM resemble closely the outputs of DCA, enabling the assessment of various product configurations and their impacts on customer preferences.

Beyond the preference modeling, the approach of network analysis provides plentiful opportunities in engineering design research. The use of network analysis implicitly carries assumptions about dependencies in data. Depending on the purpose of the analysis, the size of a network model can vary from a few nodes to hundreds or thousands of nodes containing a diverse set of products. Nevertheless, the network analysis results could be sensitive to the issue of missing data and influenced by how links are defined [39]. Another limitation for ERGM is the degeneracy problem in model estimation [38]. This may occur when a model fits the data poorly and cause the Markov chain to move towards an extreme graph of all or no edges.

MNA is not immune from practical issues when implemented to preference analysis in the context of product design. As this paper is focused on developing the conceptual framework of the proposed approach, the case study needs to be enriched by introducing more structural effects, and the developed ERGMs will be assessed by computing a simulation-based goodness-of-fit metric. In addition, the current MCPN application will be extended to incorporate other types of relations, e.g., directed association links for products and purchase preference links for customers. To calibrate and validate the simulated social network, we will conduct a small scale empirical study to assess relevant structures of customer social influences in this context. The benefits of MNA will be validated by comparing results to those from DCA, including relative quality measures on model fit (e.g. BIC) and predictive accuracy on the hold-out sample (e.g. hit rate).

While this paper focuses on modeling customer-product relations, the long-term goal of our research is to predict customer preference decisions as functions of product and customer attributes using network modeling. Such preference models will be integrated into the design optimization framework to support engineering decision making as what have been presented previously in literature.

ACKNOWLEDGMENTS

We thank National Science Foundation (CMMI-1436658), Army Research Laboratory (W911NF-09-2-0053), and Ford-Northwestern Alliance for the financial support. Additional thanks are given to Ford Corporate Economics & Global Consumer Insight group for providing the China market data.

REFERENCES

[1] Ulrich, K. T., 2003, Product design and development, Tata McGraw-Hill Education.

[2] Ben-Akiva, M. E., and Lerman, S. R., 1985, Discrete choice analysis: theory and application to travel demand, MIT press.

[3] Train, K. E., 2009, Discrete choice methods with simulation, Cambridge university press.

[4] Shiau, C.-S. N., and Michalek, J. J., 2009, "Should designers worry about market systems?," Journal of Mechanical Design, 131(1), p. 011011.

[5] Frischknecht, B. D., Whitefoot, K., and Papalambros, P. Y., 2010, "On the suitability of econometric demand models in

design for market systems," Journal of Mechanical Design, 132(12), p. 121007.
[6] Morrow, W. R., Long, M., and MacDonald, E. F., 2014, "Market-System Design Optimization With Consider-Then-Choose Models," Journal of Mechanical Design, 136(3), p. 031003.
[7] Chen, W., Hoyle, C., and Wassenaar, H. J., 2013, Decision-based design: Integrating consumer preferences into engineering design, Springer.
[8] Trist, E., 1981, "The evolution of socio-technical systems," Occasional paper, 2, p. 1981.
[9] Wasserman, S., and Faust, K., 1994, Social network analysis: Methods and applications, Cambridge university press.
[10] Monge, P. R., and Contractor, N. S., 2003, Theories of communication networks, Oxford University Press.
[11] Contractor, N., Monge, P., and Leonardi, P. M., 2011, "Network Theory| Multidimensional Networks and the Dynamics of Sociomateriality: Bringing Technology Inside the Network," International Journal of Communication, 5, p. 39.
[12] Rice, R. E., and Aydin, C., 1991, "Attitudes toward new organizational technology: Network proximity as a mechanism for social information processing," Administrative science quarterly, pp. 219-244.
[13] Lusher, D., Koskinen, J., and Robins, G., 2012, Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications, Cambridge University Press.
[14] Robins, G., Pattison, P., Kalish, Y., and Lusher, D., 2007, "An introduction to exponential random graph (p*) models for social networks," Social networks, 29(2), pp. 173-191.
[15] Wang, P., Robins, G., Pattison, P., and Lazega, E., 2013, "Exponential random graph models for multilevel networks," Social Networks, 35(1), pp. 96-115.
[16] Sosa, M. E., Eppinger, S. D., and Rowles, C. M., 2007, "A network approach to define modularity of components in complex products," Journal of mechanical design, 129(11), pp. 1118-1129.
[17] Fan, B., Qi, G., Hu, X., and Yu, T., 2013, "A network methodology for structure-oriented modular product platform planning," J Intell Manuf, pp. 1-18.
[18] Sosa, M., Mihm, J., and Browning, T., 2011, "Degree distribution and quality in complex engineered systems," Journal of Mechanical Design, 133(10), p. 101008.
[19] Cormier, P., Devendorf, E., and Lewis, K., "Optimal Process Architectures for Distributed Design Using a Social Network Model," Proc. ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 485-495.
[20] Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M., 2012, "Mine your own business: Market-structure surveillance through text mining," Marketing Science, 31(3), pp. 521-543.
[21] Aral, S., and Walker, D., 2011, "Creating social contagion through viral product design: A randomized trial of peer influence in networks," Management Science, 57(9), pp. 1623-1639.
[22] Narayan, V., Rao, V. R., and Saunders, C., 2011, "How peer influence affects attribute preferences: A Bayesian updating mechanism," Marketing Science, 30(2), pp. 368-384.
[23] He, L., Wang, M., Chen, W., and Conzelmann, G., 2014, "Incorporating social impact on new product adoption in choice modeling: A case study in green vehicles," Transportation Research Part D: Transport and Environment, 32(0), pp. 421-434.
[24] Delre, S. A., Jager, W., and Janssen, M. A., 2007, "Diffusion dynamics in small-world networks with heterogeneous consumers," Computational and Mathematical Organization Theory, 13(2), pp. 185-202.
[25] Watts, D. J., and Strogatz, S. H., 1998, "Collective dynamics of ‘small-world’networks," nature, 393(6684), pp. 440-442.
[26] Newman, M. E., and Girvan, M., 2004, "Finding and evaluating community structure in networks," Physical review E, 69(2), p. 026113.
[27] Corominas-Murtra, B., Goñi, J., Solé, R. V., and Rodríguez-Caso, C., 2013, "On the origins of hierarchy in complex networks," Proceedings of the National Academy of Sciences, 110(33), pp. 13316-13321.
[28] Page, L., Brin, S., Motwani, R., and Winograd, T., 1999, "The PageRank citation ranking: Bringing order to the web."
[29] Freeman, L. C., 1979, "Centrality in social networks conceptual clarification," Social networks, 1(3), pp. 215-239.
[30] De Vries, H., 1998, "Finding a dominance order most consistent with a linear hierarchy: a new procedure and review," Animal Behaviour, 55(4), pp. 827-843.
[31] Frank, O., and Strauss, D., 1986, "Markov graphs," Journal of the american Statistical association, 81(395), pp. 832-842.
[32] Geyer, C. J., and Thompson, E. A., 1992, "Constrained Monte Carlo maximum likelihood for dependent data," Journal of the Royal Statistical Society. Series B (Methodological), pp. 657-699.
[33] Urberg, K. A., 1992, "Locus of peer influence: Social crowd and best friend," Journal of Youth and Adolescence, 21(4), pp. 439-450.
[34] McPherson, M., Smith-Lovin, L., and Cook, J. M., 2001, "Birds of a feather: Homophily in social networks," Annual review of sociology, pp. 415-444.
[35] Cranmer, S. J., and Desmarais, B. A., 2011, "Inferential network analysis with exponential random graph models," Political Analysis, 19(1), pp. 66-86.
[36] Tan, P.-N., Kumar, V., and Srivastava, J., 2004, "Selecting the right objective measure for association analysis," Information Systems, 29(4), pp. 293-313.
[37] Gower, J. C., 1971, "A general coefficient of similarity and some of its properties," Biometrics, pp. 857-871.
[38] Snijders, T. A., 2002, "Markov chain Monte Carlo estimation of exponential random graph models," Journal of Social Structure, 3(2), pp. 1-40.
[39] Kossinets, G., and Watts, D. J., 2006, "Empirical analysis of an evolving social network," Science, 311(5757), pp. 88-90.

