2023 Conference on Systems Engineering Research

# Product Competition Analysis for Engineering Design: A Network Mining Approach

Yinshuang Xiao[a], Yaxin Cui[b], Michael T Cardone[a], Wei Chen[b], Zhenghui Sha[a*]

[a]*The University of Texas at Austin, Austin, TX 7870, USA*
[b]*Northwestern University, Evanston, IL 60208, USA*

**Abstract**

Gaining a deep insight into the factors that influence product competition is essential for a company to maintain its competitiveness in the market. While many studies have been conducted on competition analysis of various products, existing work often has oversight of market heterogeneity. This makes the analysis of product competition less accurate, which could significantly influence many downstream product design decisions. To address this issue, this paper presents a network mining approach to support product competition analysis for engineering design. The approach investigates product competition (represented by co-consideration relations) networks at three different levels, including macro (competition within the entire market), meso (competitions happening between a small group of products), and micro (competitiveness of individual products) levels. In this approach, we first develop a network motif-based representation of individual products' competitiveness. Then we use the Exponential Random Graph Model (ERGM) to study how the inclusion of such competitiveness measurement would influence products' co-consideration relations and improve the model's goodness-of-fit. This network mining approach is demonstrated in a case study on the household vacuum cleaner market, where heterogeneous customer preferences are pervasive. A multi-level network analysis of product competition provides a new way to quantify the competitiveness of a product in a heterogeneous market. It also helps quantify the importance of different competitive roles (e.g., competition within a brand or across brands) in forming co-consideration relations in the market.



*Keywords*: Network-based analysis; Data mining; Heterogeneous market systems; Market competition.

## 1. Introduction

The competitiveness of a company is the result of a combination of external and internal factors. External factors include 1) the inherent characteristics of a product market, such as its size associated with the volume of customer demand and market differentiation determined by diverse customer preferences; 2) its competitive environment

* Corresponding author. Tel.: +1-512-471-1818.
*E-mail address:* zsha@austin.utexas.edu

shaped by all market participants and stakeholders. Internal factors involve a company's organizational forms, product strategies, and the speed of its response to changing technologies and market opportunities. For example, when a new technique is introduced, a competitive company can often rapidly master it to launch new products or upgrade existing products[1]. To maintain a competitive position in the market in the long run, competition analysis is important for a firm to gain a thorough understanding of both the external and internal factors that influence its competitiveness. One external competition analysis example is to investigate the competitive environment of a market, such as studying customer preferences[2] of its representative products and typical competition patterns (e.g., how products compete between brands). An internal competition analysis example is for a company to generate a better understanding of the market positions of its own products, such as the market share of the most popular product or the one that always competes against other brands.

In recent decades, competition analysis of product markets has received significant attention. In particular, researchers in the market science domain have contributed rich findings and analysis approaches[1,3]. For example, Karuna determined the competition of a product market in three dimensions: product substitutability, market size, and entry costs. Based on this determination, he demonstrated that companies offer stronger managerial incentives when industry competition is more intense[4]. In another instance, Bustamante and Donangelo explored the interrelation between the competitive environment in which firms operate and their exposure to systematic risk[5]. More recently, researchers from engineering design communities utilized product market competition analysis to better understand the needs in engineering design. Wang et al. focused on product design for uncertain market systems[6]. They proposed an agent-based approach to help firms make competitive product design and pricing decisions to face possible reactions from market players in the short and long runs. Yip et al. investigated the possibility of using a subset of competing products or composite products to replace a large set of competing products. They found that optimal product design decision is independent of any information about competitors when customer preferences are homogeneous, but this is not valid when customer preferences are heterogeneous[7]. Wang and Chen et al. proposed using customer preference data to build competition networks. Then, various network-based competition analyses (e.g., the evolution of product competitions) were generated, which were demonstrated using the vehicle market as case studies[8–10].

The studies mentioned above primarily focused on homogeneous market analysis or a market with heterogeneous customer preferences studied by statistical models such as random-coefficients logit models[7]. In market science and economics, a perfectly heterogeneous market denotes that each small segment of demand is satisfied by just one unique segment of supply[11]. In this study, the product market is generated by customer preference data[2]. As shown in Fig. 1, the market is constructed by $M$ unique products, all of which are stated by $N$ customers through a survey study (each customer stated his/her considered products and the final purchased product). Herein, in this study, we define the *heterogeneity of a product market* as the extent to which the preferences of customers vary across the different products in the market, and we propose to use the ratio $r_h$ to measure it, as illustrated in Equation (1),

$$r_h = \frac{M}{N}. \tag{1}$$

A larger ratio indicates that customers' preferences are more scattered, resulting in a more heterogeneous market. For example, the 2013 new car buyer survey data employed by the referred study[9] contains around $M = 400$ unique vehicle models preferred by about $N = 50{,}000$ new car buyers, yielding a ratio of 0.008. In this study, we utilize the US household vacuum cleaner market data collected from our previous survey study[12]. This dataset includes 945 customers and 612 unique vacuum cleaner models; thereby, the ratio is 0.65. As a result, the household vacuum cleaner market is much more heterogeneous than the vehicle market case because, for each customer in the vacuum cleaner market, there are more product options per customer.

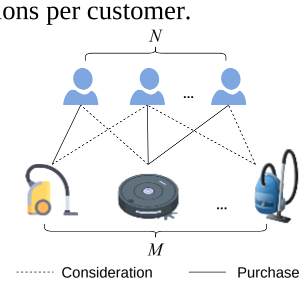



Fig. 1. Illustration of product market.

Even though recent efforts have enriched product market competition analysis through customer preference modeling and an in-depth understanding of the impacts of the market environment on a company's operation, a fundamental research gap remains in learning the competition relationships in a highly heterogeneous market. One major challenge here is the characterization and quantification of different types of competition (inter-brand and intra-brand competitions) due to market heterogeneity. To address this challenge, this study developed a multi-level network mining approach to studying product competitions from macro (competition within the entire market), meso (competitions happening between a small group of products), and micro (competitiveness of individual products) network levels. This approach integrates network motifs into ERGM to represent micro-level product competitiveness and measure the influence of product competitiveness on the customers' consideration and choice decisions. We demonstrate the utility of the approach with a case study on the household vacuum cleaner market.

## 2. Knowledge Background

This study employs two network analysis techniques based on network motif theory and the exponential random graph model (ERGM). The technical background and the associated technical details are introduced below.

### *2.1. Network Motif*

Network motifs[13] are underlying non-random subgraphs within complex networks. They can be classified as directed or undirected and can also be categorized by the number of nodes they consist of. The most common statistic to assess the significance of a network motif in a complex network is motif *Z-score*. Given a graph $G$ and an $n$-size motif $G'$, the frequency of $G'$ in $G$ is the number of times that $G'$ appeared in $G$, which is denoted by $F_G(G')$. Then, considering an ensemble of random graphs corresponding to the null model of $G$ be $\omega(G)$, $R(G)$ is a set that includes $K$ randomized networks, all of which are from $\omega(G)$. Accordingly, the *Z-score* is defined as

$$Z_G(G') = \frac{F_G(G') - \mu_R(G')}{\sigma_R(G')}\,. \tag{2}$$

Another common metric for the evaluation of significant network motifs is the *P-value*. It indicates the probability of $F_R(G') > F_G(G')$, where $F_R(G')$ is the frequency of $G'$ in random network $R$. *P-value* can be calculated by

$$P_G(G') = \frac{1}{K}\sum_{j=1}^{K} \delta(F_R(G') > F_G(G')), \tag{3}$$

where $K$ represents the total number of considered random networks, and $j$ is the index of each random network. $\delta$ equals 1 when $F_R(G') > F_G(G')$, and 0 otherwise. In general, one motif pattern is significant if the *P-value* is smaller than a typical threshold, commonly 0.01 or 0.05.

### *2.2. Exponential random graph models*

ERGM is a family of statistical inference models for network data analysis[14]. The basic assumption of these models is that an observed network $\boldsymbol{y}$ is one specific realization from a set of possible random networks $\mathbb{Y}$, and its probability model follows the distribution in Equation (4).

$$\Pr(\boldsymbol{Y} = \boldsymbol{y}) = \frac{\exp\left(\boldsymbol{\theta}^T \cdot \boldsymbol{g}(\boldsymbol{y})\right)}{\kappa(\boldsymbol{\theta})}, \boldsymbol{y} \in \mathbb{Y}. \tag{4}$$

where $\boldsymbol{g}(\boldsymbol{y})$ is a vector of the model statistics defining various network structures that can incorporate either nodal attributes or edge attributes, $\boldsymbol{\theta}$ is a vector of model coefficients associated with $\boldsymbol{g}(\boldsymbol{y})$, and $\kappa(\theta)$ is a normalizing constant to make sure Equation (4) generates a probability value in $[0, 1]$.

## 3. Methodology

An overview of the proposed multi-level network mining approach is presented in Fig. 2. From top to bottom, in **Layer One**, we collect customer preference data, including the product alternatives they consider (i.e., choice sets) and the final choices (assuming each customer chooses only one product from his/her choice set). In **Layer Two**, our objective is to investigate the overall characteristics of the competition within a product market by analyzing two-stage competition networks at the macro level. Referring to existing studies[2,12], we first build two macro-level unidimensional networks, a co-consideration network, and a choice network, to model competition relationships using the customer preference data. The nodes in both networks represent the unique product models considered by customers. In the co-consideration network, the links are undirected and represent co-consideration relations between two products. In the choice network, the links are directed, denoting two products being co-considered, but the direction points to the one that was purchased. Once the networks are constructed, various network metrics, such as average node degree, global cluster coefficient, and network density, are adopted for network analysis to draw insights into the market competition[15].

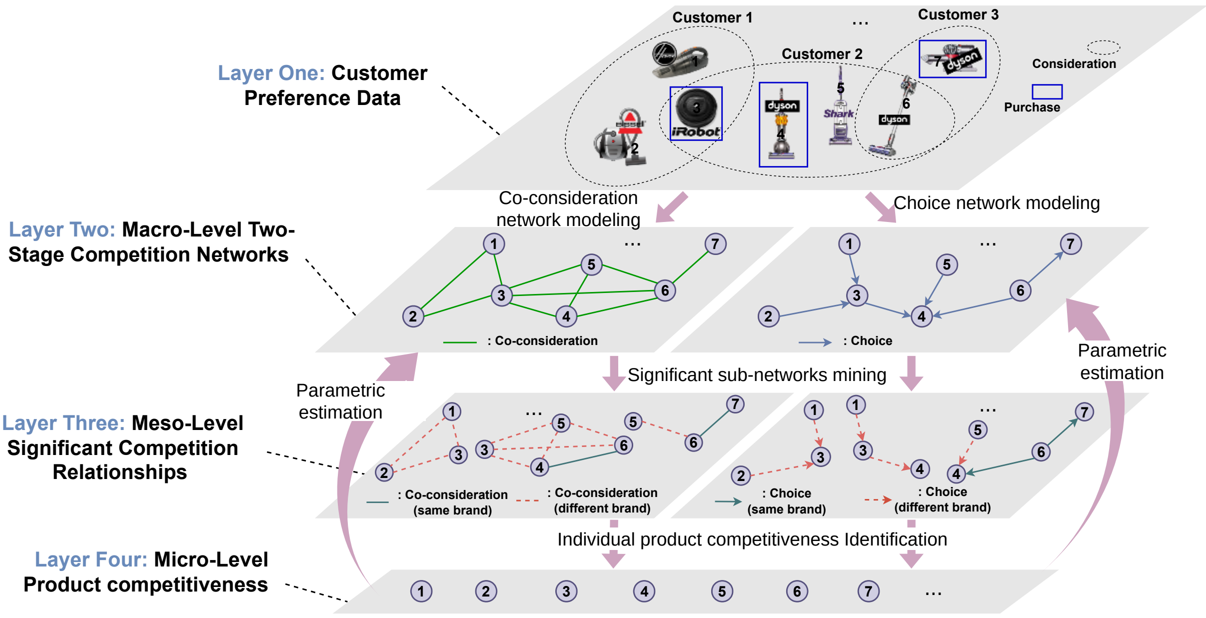



Fig. 2. The proposed multi-level network mining approach for product market competition analysis.

In **Layer Three**, we aim to identify the significant meso-level competition patterns. In this study, we differentiate the competition between two products across brands and those within a brand. This is a fair assumption because, from the business perspective, these two types of competition can provide different directions for a company. For instance, fierce cross-brand competition serves as a reminder for businesses to create products that set themselves apart from their rivals, whereas intense competition within a brand indicates that businesses must modify their product lines to prevent serious cannibalization[16]. Therefore, we label the links in both networks into two types based on whether or not two connected products belong to the same brand. As shown in Fig. 2, a dashed link represents two products from different brands, and a solid link represents the same brand. Next, a network motif mining tool, FANMOD[17], is employed to enumerate the significant competition motifs in both networks. The benefits of identifying these significant competition motifs are: 1) the local pattern of competition relationships of each brand can be discovered by analyzing the topologies of the motifs, and 2) the insights into the competitiveness of each product can be assessed by analyzing the roles a product played in the significant network motifs.

In **Layer Four**, we introduce the concept of *node role*[18,19], which is described as the role that a node plays in a network motif structure, e.g., the center of a star network structure. After obtaining node roles, we describe the competitiveness of each product based on the number of times it is involved in a role. Then the competitiveness, treated as one product attribute, is used in ERGM modeling of the macro-level competition networks. A more detailed discussion based on a case study is given in Section 4.

## 4. Case Study

In this section, we use the US household vacuum cleaner market as a case study to demonstrate our proposed multi-level data mining approach.

### *4.1. Data source*

The dataset used in this study is drawn from our prior survey study[12] conducted in 2021 by Cint – a company providing digital survey solutions. The dataset includes 945 customers' responses to 612 unique household vacuum cleaner models. The dataset covers a wide range of customer-specific and vacuum cleaner-specific attributes, such as customer demographics, technical features of vacuum cleaners, customer social relationships, and so on. The respondents were required to list their considered vacuum cleaners and the ones they ultimately bought.

### *4.2. Macro-level network analysis*

Following the methodology presented in Section 3, we first construct the unidimensional co-consideration and choice networks based on the survey data of the 945 customers. Given the limited data source, whenever a customer co-considers two vacuum cleaners, a link will be formed between these two models, and the link weight indicates the number of times that the two products are co-considered. The co-consideration network includes 612 nodes and 2058 undirected links. The choice network includes 612 nodes and 1187 directed links.

In this study, we focus on the top-ten dominant brands in the market. These brands are identified by ranking the frequencies that the respondents considered and purchased the products. In other words, in the original competition networks, we only keep the links that connect the products of the top-ten brands. The visualizations of the downscaled co-consideration and choice networks are shown in Fig. 3. As a result, the heterogeneity of this top-ten market decreases from 0.65 to 0.45. However, this value is still much higher than the 2013 new car buyer survey data (0.008), and the market is considered highly heterogeneous.

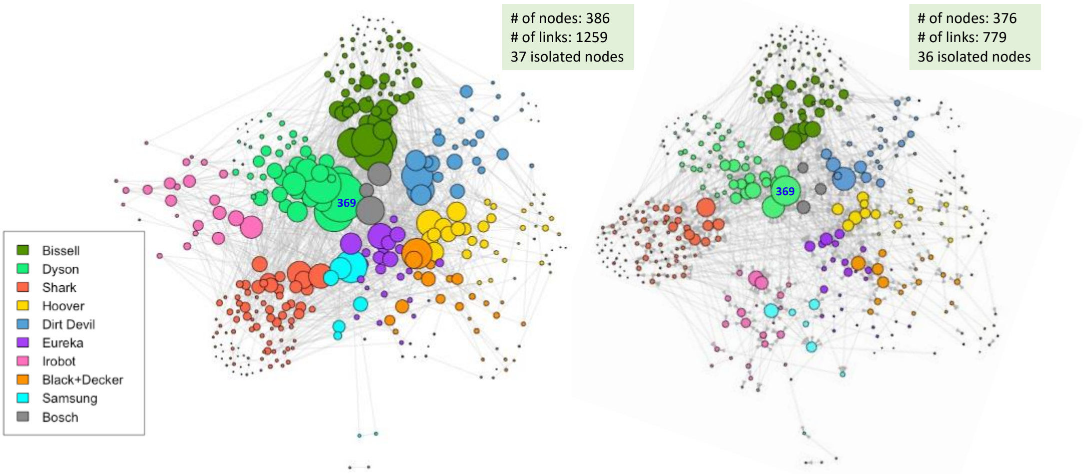
Next, we analyze the vacuum cleaner market competition in both consideration and choice stages by network metrics, and the results are summarized in Table 1. The network density measures competition intensity. The average unweighted degree reveals the average number of products a product competes against, and the average weighted degree shows the average number of competition relations a product involves. Taking the co-consideration network as an example, a vacuum cleaner competes against 6.5 other vacuum cleaners on average and is co-considered 6.85 times on average. The average local cluster coefficient measures how likely two competing products both compete with the same product on average. For instance, in the co-consideration network, the average probability that two competing vacuum cleaners compete with one common vacuum cleaner is 0.433, which is higher than that of the choice network, 0.068. This denotes that the clustering of competitions is more likely to occur at the consideration stage than at the choice stage. In short, these measurements provide us with an overview of the competitive environment at the market level, including the overall competition intensity, the clustering level of competition, as well as the average competition intensity at the product level.

Table 1. Network metrics to quantify the properties of macro-level competition networks

| | Network density ($\rho$) | Average unweighted ($\bar{d}$), weighted ($\bar{d}'$) degree | Average local cluster coefficient ($\bar{c}$) |
|---|---|---|---|
| Co-consideration network | 0.017 | 6.52, 6.85 | 0.433 |
| Choice network | 0.006 | 4.14, 4.20 | 0.068 |

### *4.3. Meso-level significant competition network motif identification and interpretation*

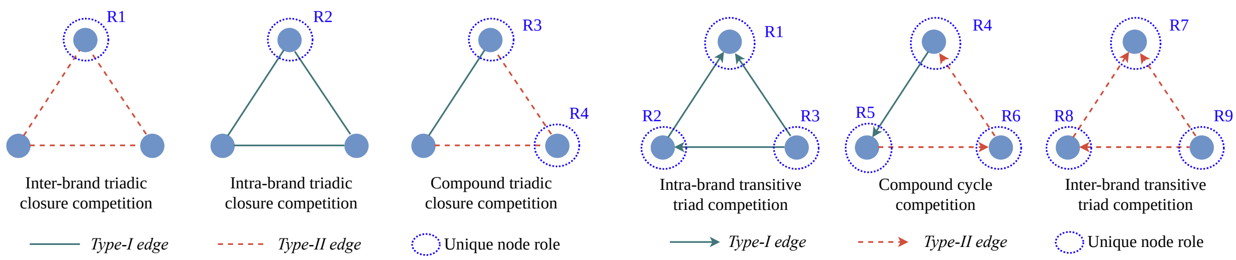
As illustrated in Section 3, we first label the edges in the competition networks into two types: *type-I edge* indicates that two vacuum cleaners share the same brand, and *type-II edge* refers to the different brand types. Given the essential role of triad census in networks science[20], we concentrate on size-3 sub-networks. Then, the network motif mining tool, FANMOD, is adopted to identify the most significant size-3 motifs of competition. As shown in Fig. 4, each motif represents distinct competition relationships between brands (inter-brand) and within a brand (intra-brand), and they are named by their edge types and topological characteristics.




(a) Co-consideration Network (b) Choice Network

Fig. 3. Competition networks of top-ten household vacuum cleaner brands. The legend is ranked based on the popularity of each brand in the consideration stage, and the size of nodes is proportional to the node degree. For example, product 369, Dyson Ball Multi-floor 2, shows the largest node size in both networks. Its unweighted degree and weighted degree in the co-consideration network are 38 and 47, indicating that it is co-considered with 38 models for a total of 47 times. Its unweighted degree, weighted degree (sum of weighted in- and out-degree), and weighted in-degree in the choice network are 31, 33, and 19, respectively, indicating that it competed with 31 models 33 times and purchased 19 times.

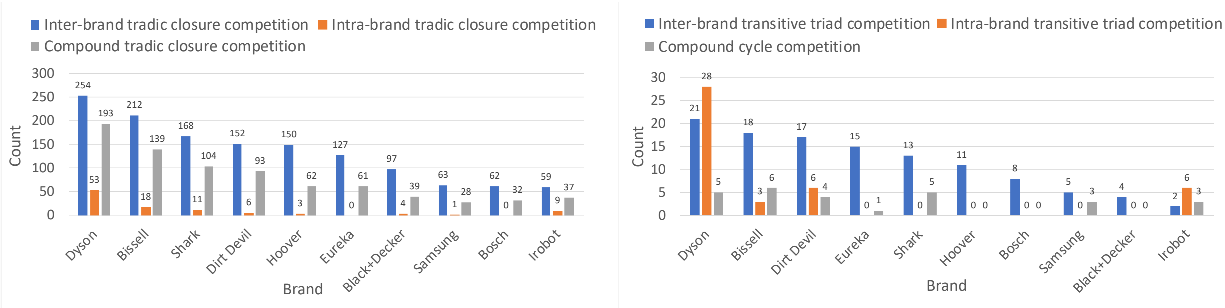
These competition motifs, which are discovered to be significant, allow us to perform two types of analyses. The first is micro-level *node role* identification, which is presented in Section 4.4. The second is meso-level brand competition quantification, and an example is presented in Fig. 5. The numbers of various competition motifs in which each brand participates are listed in Fig. 5. These analysis results provide insights into competitive trends across brands at both the consideration and choice stages. It is found that in the consideration stage, the inter-brand triadic closure competition is the dominant local competition in this vacuum cleaner market. Dyson is the most competitive brand, as evidenced by its more frequent participation in all three of these competition motifs than other brands. In the choice stage, except for Dyson and iRobot, which are more frequently involved in the intra-brand transitive triad competition, the inter-brand transitive triad competition is the most frequently-occurring competition for most brands. Another insight we can derive from Fig. 5 is the competition types that one particular brand participated in at two stages. For example, Dyson is more often considered alongside other brands during the consideration stage, whereas in the choice stage, the competition more frequently happens within the Dyson family.




(a) Significant size-3 competition motifs and corresponding node roles in co-consideration network

(b) Significant size-3 competition motifs and corresponding node roles in choice network

Fig. 4 significant size-3 competition motifs and corresponding node roles in two-stage competition networks.




(a) Co-consideration network (b) Choice network

Fig. 5 Number of times that each brand involves in each type of competition motif. (a) is ranked by inter-brand triadic closure competition; (b) is ranked by inter-brand transitive triad competition.

### *4.4. Micro-level product competitiveness*

We define different types of *node roles* in the co-consideration network and choice network, respectively, based on the position where a node locates in the motifs, highlighted by dot circles in Fig. 4. For the motifs in the co-consideration network, there are four node roles, each of which represents a distinct competitive position. For example, role R1 delineates the competition when one product competes with two products from the other two different brands, whereas R2 represents the competition among products with the same brand. In the choice network, the topology becomes more complex because of the existence of link direction. So, there is a total of nine distinct node roles, depending on the position of each node in the motifs, as shown in Fig. 4 (b).

After generating *node roles*, we define the competitiveness of each vacuum cleaner model as

$$C_{ij} = \frac{n_i^{R_j}}{d_i}, i \in M, j \in [1,9], \tag{5}$$

where $n_i^{R_j}$ indicates the number of times that product $i$ is involved in the role $R_j$, and $d_i$ is the network degree of product $i$ to normalize $n_i^{R_j}$. Taking Hoover Powerdrive as an example, it is involved in R1 four times, R2 zero times, R3 three times, and R4 two times in the co-consideration network. Its network degree is 14; therefore, its competitiveness vector is [0.29, 0, 0.21, 0.14]. This indicates that in three-way competitions, it more frequently competes with products from distinct brands rather than within the Hoover product family.

### *4.5. The impact of product competitiveness on the competition network formation*

To study the impact of micro-level product competitiveness on the macro-level competition formations, we employ ERGM to study the estimates of node attributes. The considered attributes are categorized into four types. The first type is the network configuration. Given that ERGM with complex network configurations (e.g., star-type interdependence and triangle-type interdependence[9]) suffers from a model degeneracy issue[21], i.e., failing to generate statistically significant representations of the observed networks in Fig. 3, only edges, equivalent to the intercepts of the regular regression model, are considered. The second type is the baseline effect of vacuum cleaner attributes such as suction power and price[9]. The third type is the homophily effect (i.e., matching and difference) of vacuum cleaner attributes, such as price difference between co-considered vacuum cleaners[9]. The fourth type is the value of the competitiveness vector obtained in Section 4.4. In Equation (4), all these four types of attributes are regarded as the model statistics in modeling. Then, the Markov Chain Monte Carlo (MCMC) algorithm is adopted to estimate the corresponding coefficient vector $\boldsymbol{\theta}$ to maximize the likelihood of observed network structures aggregately. Finally, insights into the importance of each attribute are generated by assessing the *p-values* as well as the sign and magnitude of the estimated coefficients[22]. In this study, we compare the estimation results between those with and without product competitiveness features and evaluate the model's goodness-of-fit by the spectral goodness-of-fit (SGOF) metric[23], as shown in Equation (6).

$$SGOF = 1 - \frac{E\bar{S}D_{obs,fitted}}{E\bar{S}D_{obs,null}} \tag{6}$$

where $E\bar{S}D_{obs,fitted}$ and $E\bar{S}D_{obs,null}$ represent the average Euclidean spectral distance under the fitted model and null model, respectively. SGOF measures the improvement of using a fitted model to describe the observed macro-level network over the null model, and an SGOF of 1 indicates a perfect description[23]. We present the estimation results of the co-consideration network only in Table 2 as an example.

Table 2. ERGM-based estimation results of the co-consideration network

| Input variables | ERGM (without competitiveness representation) | | ERGM (with competitiveness representation) | |
|---|---|---|---|---|
| | **Estimate coefficient** | **Std. Error** | **Estimate coefficient** | **Std. Error** |
| ***Network configurations*** | | | | |
| Edges/Intercept | -3.62 *** | 0.20 | -5.77*** | 0.21 |
| ***Baseline effects of vacuum cleaner attributes*** | | | | |
| Vacuum type[a] (handheld) | -0.39 *** | 0.22 | -0.51*** | 0.11 |
| Vacuum type (robotic) | 0.03 | 0.10 | 0.06 | 0.11 |
| Vacuum type (stick) | -0.41 *** | 0.08 | -0.44*** | 0.08 |
| Vacuum type (upright) | -0.42 *** | 0.08 | -0.27*** | 0.08 |
| Suitable for pet hair (binary) | -0.21 *** | 0.06 | -0.14* | 0.06 |
| HEPA filter (binary) | -0.15 ** | 0.05 | -0.10 | 0.05 |
| Price (continues) | 1.21 *** | 0.17 | 0.76*** | 0.18 |
| Suction power (continues) | 0.58 *** | 0.09 | 0.25** | 0.09 |
| Warranty (continues) | -0.29 ** | 0.10 | -0.24* | 0.11 |
| ***Homophily effects of attribute matching and difference*** | | | | |
| Vacuum type matching | 0.84 *** | 0.06 | 0.89*** | 0.07 |
| Suitable for pet hair matching | 0.12 | 0.07 | 0.10 | 0.07 |
| HEPA filter matching | 0.02 | 0.06 | 0.05 | 0.06 |
| Price difference | -2.41 *** | 0.24 | -2.40*** | 0.25 |
| Suction difference | -0.18 | 0.12 | -0.09 | 0.12 |
| Warranty difference | -0.50 *** | 0.14 | -0.53*** | 0.14 |
| ***Vacuum cleaner competitiveness (continuous)*** | | | | |
| R1 | - | - | 2.21*** | 0.10 |
| R2 | - | - | 1.78*** | 0.12 |
| R3 | - | - | 1.38*** | 0.09 |
| R4 | - | - | 1.13*** | 0.11 |
| ***Model performance*** | | | | |
| Bayesian Information Criterion (BIC) | 12430 | | 11248 | |
| Mean SGOF (5$^{th}$, 95$^{th}$ quantiles) | 0.11 (0.06, 0.16) | | 0.77 (0.71, 0.82) | |

***: *p-value* < 1e-04, **: *p-value* < 0.001, *: *p-value* < 0.01

[a]: the baseline of vacuum type is the canister

All the continuous variables are normalized by max-min normalization

The estimated coefficients in the two ERGMs have identical signs and similar magnitudes. For example, the estimated coefficient of the suction power is 0.58. The positive sign and its level of statistical significance indicate that the vacuum cleaners, both with higher suction power, are about 1.8 (i.e., $e^{0.58}$) times more likely to be co-considered together than those with low suction power. In the group of homophily effects, for example, the negative sign of price difference indicates that the vacuum cleaners with similar prices are more likely to compete against each other. Another example is the baseline effect of price. Its positive sign indicates that the vacuum cleaners with higher prices are 3.4 (i.e., $e^{1.21}$) times more likely to be co-considered than those with low prices, suggesting that vacuum cleaner customers are not price-sensitive and won't forego comparing a vacuum cleaner with another one because of its price.

In the second model, where we include the competitiveness measurement, we found additional interesting results. First, the signs of the four types of competitiveness R1 – R4 are all positive and statistically significant. This implies that competitive products (i.e., being more frequently involved in the roles of R1, R2, R3, and R4) are always more likely to be co-considered. Additionally, the larger magnitude of the R1 coefficient shows that the role of R1 contributes more to the formation of the co-consideration links than the other three node roles. This may imply that if a producer of vacuum cleaners wants to consolidate its competitive position in the market, it would be better to get involved in inter-brand triadic closure competition more frequently.

As to the model performance, the lower BIC value and the higher SGOF value demonstrate that ERGM with competitiveness measurements reproduces the observed competition networks better than the one without competitiveness measurements. Note that such improved goodness-of-fit value may imply overfitting, so the model would be less capable of predicting new networks with unseen data. It also leads to a question of whether the second model is prone to produce a causality dilemma, i.e., using the attributes produced by the original macro-level networks to estimate the same networks. However, we believe this dilemma does not affect the interpretation of the relative importance of each competitiveness attribute. As a result, in this study, the proposed estimation model serves as a supplementary model of ERGM without considering complex network configurations to aid in understanding the competitive roles that a product plays in the market. In our future work, more efforts will be spent to address the causality dilemma issue to make the model with adequate predictive power.

## 5. Conclusion

This paper proposes a multi-level network mining approach to support competition analysis of heterogeneous product markets. This approach starts with transferring customer preference data into two-stage (consideration and choice) competition networks, followed by a macro-level competition study using descriptive network analysis. Then, the meso-level critical competition motifs (or subgraphs) are identified and explored. Next, we develop a network motif-based representation of micro-level products' competitiveness. Finally, the impact of those identified product competitiveness features on the formation of competition networks is studied using ERGM. We demonstrated the approach with a case study on the household vacuum cleaner and obtained insights into the market competition for the top-ten brands. However, due to the potential causality dilemma issue, the ERGM model with competitiveness features included is suggested to be used for interpretation only at this stage. More studies are required to validate its utility in link prediction, which is the focus of our future work.

## Acknowledgments

The authors acknowledge collaborators Neelam Modi, Jonathan Haris Januar, and Gracia Cosenza for their assistance in data collection, data processing, and the inputs provided during research meetings. We also greatly acknowledge the funding support from NSF CMMI #2005661 and #2203080.

## References

1. Sanchez, R. Strategic flexibility in product competition. *Strategy. Manag. J.* **16**, 135–159 (1995).
2. Cui, Y. *et al.* A Weighted Statistical Network Modeling Approach to Product Competition Analysis. *Complexity* **2022**, 1–16 (2022).
3. Spiegler, R. Choice Complexity and Market Competition. *Annu. Rev. Econom.* **8**, 1–25 (2016).
4. Karuna, C. Industry product market competition and managerial incentives. *J. Account. Econ.* **43**, 275–297 (2007).
5. Bustamante, M. C. & Donangelo, A. Product Market Competition and Industry Returns. *Rev. Financ. Stud.* **30**, 4216–4266 (2017).
6. Wang, Z., Azarm, S. & Kannan, P. K. Strategic Design Decisions for Uncertain Market Systems Using an Agent Based Approach. *J. Mech. Des.* **133**, 1–11 (2011).
7. Yip, A. H. C., Michalek, J. J. & Whitefoot, K. S. Implications of Competitor Representation for Profit-Maximizing Design. *J. Mech. Des.* **144**, 1–7 (2022).
8. Wang, M. *et al.* Predicting product co-consideration and market competitions for technology-driven product design: a network-based approach. *Des. Sci.* **4**, e9 (2018).
9. Sha, Z. *et al.* A Network-Based Approach to Modeling and Predicting Product Coconsideration Relations. *Complexity* **2018**, 1–14 (2018).
10. Xie, J. *et al.* Data-Driven Dynamic Network Modeling for Analyzing the Evolution of Product Competitions. *J. Mech. Des.* **142**, 1–14 (2020).
11. Alderson, W. The Heterogeneous Market and the Organized Behavior System. in *A Twenty-First Century Guide to Aldersonian Marketing Thought* 189–215 (Springer-Verlag). doi:10.1007/0-387-28181-9_13.
12. Xiao, Y. *et al.* Information Retrieval and Survey Design for Two-Stage Customer Preference Modeling. *Proc. Des. Soc.* **2**, 811–820 (2022).

13. Milo, R. *et al.* Network Motifs: Simple Building Blocks of Complex Networks. *Science (80-. ).* **298**, 824–827 (2002).
14. Harris, J. *An Introduction to Exponential Random Graph Modeling*. *An Introduction to Exponential Random Graph Modeling* (SAGE Publications, Inc., 2014). doi:10.4135/9781452270135.
15. Wang, M. *et al.* Forecasting Technological Impacts on Customers' Co-Consideration Behaviors: A Data-Driven Network Analysis Approach. in *Volume 2A: 42nd Design Automation Conference* vols 2A-2016 (American Society of Mechanical Engineers, 2016).
16. Chen, F. Studying Product Competition and Sequential Targeting Using Big Data and Machine Learning. *ProQuest Dissertations and Theses* (New York University PP - United States -- New York, 2022).
17. Wernicke, S. & Rasche, F. FANMOD: a tool for fast network motif detection. *Bioinformatics* **22**, 1152–3 (2006).
18. Scripps, J., Tan, P.-N. & Esfahanian, A.-H. Node roles and community structure in networks. in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07* 26–35 (ACM Press, 2007). doi:10.1145/1348549.1348553.
19. Henderson, K. *et al.* RolX- Structural Role Extraction & Mining in Large Graphs. in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '12* 1231 (ACM Press, 2012). doi:10.1145/2339530.2339723.
20. Ribeiro, P., Paredes, P., Silva, M. E. P., Aparicio, D. & Silva, F. A Survey on Subgraph Counting: Concepts, Algorithms, and Applications to Network Motifs and Graphlets. *ACM Comput. Surv.* **54**, 1–36 (2022).
21. Mukherjee, S. Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics. *Bernoulli* **26**, 1016–1043 (2013).
22. Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M. & Morris, M. ergm : A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *J. Stat. Softw.* **24**, 1–29 (2008).
23. Sha, Z., Wang, M., Company, F. M., Huang, Y. & Contractor, N. MODELING PRODUCT CO-CONSIDERATION RELATIONS: A COMPARATIVE STUDY OF TWO NETWORK MODELS. in *DS 87-6 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 6: Design Information and Knowledge* vol. 6 317–326 (2017).