

## Where Does My Product Stand? A Social Network Perspective on Online Product Reviews

Naga Vemprala  
The University of Texas at  
San Antonio  
[Naga.Vemprala@utsa.edu](mailto:Naga.Vemprala@utsa.edu)

Rachael Ruizhu Xiong  
The University of Texas at  
San Antonio  
[ruizhu.xiong@utsa.edu](mailto:ruizhu.xiong@utsa.edu)

Charles Zhechao Liu  
The University of Texas at  
San Antonio  
[Charles.Liu@utsa.edu](mailto:Charles.Liu@utsa.edu)

Kim-Kwang Raymond Choo  
The University of Texas at  
San Antonio  
[raymond.choo@fulbrightmail.org](mailto:raymond.choo@fulbrightmail.org)

### Abstract

*Customer reviews often include comparative comments on competing products. Adopting the "The Strength of Weak Ties" theory, we build a product social network around "strong tie" and "weak tie" entities. By performing text mining on comparative customer reviews collected from Amazon, we successfully identify strong and weak ties in a product network and compute the strength of these ties. Utilizing these network properties, we generate network graphs based on different product features and discover the underlying competitive relationships among them. In particular, our regression analysis shows that the strength of ties positively contributes to the review rating of a product and the strength of weak ties plays a more significant role than the strength of strong ties. These results will benefit vendors in online market to discover potential competitors, effectively tailor their marketing and product development efforts, and better position their products to increase profit and explore new market opportunities.*

### 1. Introduction

In a competitive market, in addition to traditional metrics such as price and quality, how to strategically position a product also plays a critical role in determining the market performance of the product. The accurate positioning of a product can enable manufacturers to maximize the product's market potential and minimize fierce competition from close competitors. However, doing so requires the manufacturers to have a thorough understanding of the strength and weakness of their products relative to

their close competitors, as well as potential new entrants that they will compete against in the future.

One of the resources to develop this insight and understand product competition is through analyzing online customer reviews [6, 8, 17], especially reviews that compare the product with its competitors. When shopping online, customers often search for, evaluate, and compare products through online customer reviews. Customer reviews reflect consumers' level of satisfaction with a product and provide valuable information about their collective evaluation on various dimensions of the product. When properly utilized, online reviews can help manufacturers target more profitable market segments to better meet customer demand and increase sales. Adopting a social network perspective, we contend that customer reviews are important instruments through which products are connected and form a competitive product network. Such a product network can reveal important competitive dynamics that are not immediately visible from sales statistics. Therefore, we propose to build a product social network by extracting comparative information among competing products from online customer reviews and use this network to examine the relative positions of products and how their network properties affect their market performance.

The construction of such a product network will provide several important insights. First, even though review ratings provide useful feedback on customer's perception of product quality, delving into customer reviews allows vendors to discover what specific product features or competitive advantages drive the success (or failure) of their products, so that they can better improve their product design [6]. Second, marketing and product development strategies have traditionally targeted direct competitors (i.e.

competing products that share similar product features and price range). However, other products in adjacent product space may possess unique features that constitute potential threat. Hence it is important to evaluate indirect competition to detect such threat or explore potential opportunities to cover additional market segments. Third, when evaluating products with distinct features, consumers may find a product social network very helpful in ranking products that are difficult to compare side by side, as such a product social network has the potential to uncover hidden and obscure competitive relationships.

To set up a theoretical foundation for forming such a product social network, we survey the literature on social network analysis (SNA) with a focus on finding constructs that can capture the relationship between competing products. One of the most widely used theories in the SNA literature is Granovetter's "The Strength of Weak Tie" theory [12]. In this theory, individuals in a social network are separated into strong ties and weak ties. The individuals that have direct interactions are defined as strong ties and those who do not have direct connections but have mutual friends are considered weak ties. Generally speaking, since individuals have easier access to their strong ties than to weak ties, they are more likely to neglect the values of weak ties and try to rely on strong ties for their tasks and needs. Despite this inertia, Granovetter argues that in many cases people may benefit more from weak ties than strong ties. This is largely due to the fact that strong ties often come from an individual's surrounding social groups, share similar background and characteristics, and have overlapping social capital. In contrast, the weak ties may possess unique information and social capital not typically available to strong ties. Thus, the heterogeneity brought by weak ties can provide valuable external information and enrich the scope of interaction. For example, evidence shows that weak ties positively affect the outcome of an individual's job search and job advances [19], reduce information asymmetry among social groups, and widen the pool of social resources when these weak ties are connected by more centralized nodes in a social network [3]. Therefore, weak ties are valuable resources that bridge people to a broader information network. The distinction and relationships between weak and strong ties closely resemble products competing in an online market, which lends them as good theoretical lens to examine the underlying product social network and the associated competition in such networks.

In addition to their similarity to strong and weak ties, products listed in the same category may exhibit varying degree of competitive relationship. A product may intensely compete with some products while

loosely overlap with the market segments covered by other products. Even among the same type of ties, the level of interaction varies substantially. Therefore, we also extend the "strength of tie" concept from the SNA literature to help us better capture the extent of competition among various products. The strength of tie indicates how strongly the nodes in a network are connected through the ties and is often measured by the closeness of relation between the ties [13] (i.e. the frequency and duration of contact). Typically, the stronger the tie, the more their social neighborhood overlaps [2] and in the context of product network, the more intense the competition and the larger the overlap of their targeted market segments.

With these social network concepts characterizing the competitive relationships among products, we will build a product social network consisting of various product nodes connected through strong and weak ties, with varying strength of ties. To determine how and whether the various products are directly or indirectly related, we will perform feature extraction and text mining on comparative online reviews collected from Amazon.com. By extracting the product feature associated with the specific comparative reviews, we will be able to assess how a product feature influences customers' perception on the relative advantages and disadvantages of the associated products. In particular, we will recover the underlying relationships among products by differentiating the strong ties, which are formed when products are directly compared and evaluated, from weak ties, which exist between products that have not been directly compared but are both connected to a third product through a strong tie.

The results of our study have several important theoretical and empirical implications. In the existing SNA literature, the fundamental entities are often individuals or organizations. To the best of our knowledge, we are among the first studies to apply social networking theory and analysis to studying the competitive relationships among products. Therefore, our new approach will contribute to the literature by providing a unique perspective on understanding the competitive dynamics in online markets. Moreover, the type and strength of ties defined in our proposed product network will provide a novel way to assess the relative strength and weakness of a product and complement common performance indicators such as sales volume or review rating. From a practical standpoint, the proposed theoretical framework and the empirical methodology adopted to build such a product network will help manufacturers better position their products and develop marketing strategies to cover the potential under-explored market segments.

In the next section, we will review the theoretical basis for applying the social network concepts to online product competition, followed by a description of our research methodology in Section 3, and a summary of empirical results generated from the analysis of a product social network in Section 4. Section 5 concludes the paper with potential limitations and directions for future research.

## 2. Theoretical Background and Hypotheses

In SNA, a tie is a form of connection between the various entities (nodes) in a network [2]. Social ties, also called interpersonal ties, describe the process of exchanging information among individuals and have been a key instrument for studying various types of networks across many different disciplines. For example,

- in marketing research, research shows that social ties affect customers' behaviors of online shopping in different stages [10];
- in economics research, some economists have found contradicting evidence against the well-known claim by Granovetter that weak ties may be more helpful than strong ties. By analyzing six million users on Facebook, researchers discover that even though weak ties provide a greater extent of outreach, the marginal value of a strong tie is still a lot more significant than that of a weak tie [11]; and
- in information system research, researchers find that students chose class groups based on their social-tie relationships on Facebook, rather than their peers' school performances [4].

In general, there are three types of social ties: strong, weak, and absent [13], with their intensity measured by the strength of these ties. Next, we examine the definition of these concepts and how they can be mapped to our product competition context.

### 2.1. Strong Ties

Granovetter originally measures social ties by the distance between two individuals. Strong ties are the closest relationship. The closeness is jointly measured by the frequency of interactions, ease of contact, and the extent to which they share similar attributes [14]. Due to these measurement criteria, it is not surprising that the exchange of information among strong ties is more likely to repeat and results in redundant information [3, 12]. Consequently, when it comes to

information seeking, the information possessed by strong ties may be useless for the individual they are connected with. We adopt this concept and extend it to examining products relationships in online markets. In our study, products that are direct competitors in a market (i.e. those that are introduced to the market during the same time period, fall into the same price range and have very similar features) and are directly compared by customers in their reviews are considered strong ties. For example, in our data sample, iPhone X and Samsung Galaxy Note 8 are direct competitors in the cellphone market and it is very common to see customers post reviews that involve a direct comparison between these two products. Therefore, we will consider iPhone X and Samsung Galaxy Note 8 to be strong ties in a product social network. Similar to strong ties in ordinary social networks, strong ties in a product network may also exhibit very similar technical specifications, and often constitute strong substitutes from a competitive perspective.

### 2.2. Weak Ties

Weak ties are another form of social ties from Granovetter's research. They are two indirect contacts that share a mutual friend [14]. In the original study, Granovetter proposes that weak ties are more effective because they are from different social groups and contain more unique information. Such a relationship can be considered "a friend of a friend". In Granovetter's research, he claims that weak ties are the essential networks for job searching [12, 14]. Much of later research shows that weak ties are more helpful for non-redundant information transmission. For example, sociology studies show that people are more likely to marry weak-tie contacts because of the unfamiliarity and attractions [20]. In our study, products that have not been directly compared by customers but have been both compared with a common third product (i.e. "mutual friend") are considered weak ties. For example, in this study, iPhone X and Nokia 6 are not direct competitors in the cellphone market and no customer review compares these two products. However, some customers compared both iPhone X and Nokia 6 with Google Pixel. So, in this case, iPhone X and Nokia 6 are considered weak ties. Weak ties in our study are not direct competitors on the whole, but may have certain product features that enable them to become potential substitutes or competitors with the focal product.

### 2.3. Strength of Tie

The strength of tie was introduced in Granovetter's seminal research "The Strength of Weak Ties" [14].

By conducting a survey on 282 professional, technical, and managerial workers, Granovetter finds that the strength of ties can be measured by the frequency of contacts. The more frequent two individuals interact, the stronger the relationship in their social network. In subsequent studies where the network is used to capture interpersonal communication, the strength of tie has often used to convey the closeness in relationship between individuals, and is measured by duration of relationship, extent of emotional support and overlapping social ties [2, 12]. In social media networks, the strength of tie is measured using social media characteristics such as shared wall posts, photos mutually tagged, normalized measure of number of common friends, or embeddedness. For example, in [1], the strength of tie is measured as the number of common friends divided by the total number of friends.

In our study, we will use the strength of tie to represent the extent of direct or indirect competition among products and derive our hypotheses based on the strong and weak ties theory.

## 2.4. Hypotheses

Once we identify the type of ties and obtain their tie strength, we will be able to use these network properties to derive some underlying relations between products and discover how these relationships can be used to assess or predict product performance. Since the review rating of a product has been widely used to gauge the consumers' satisfaction with the product, it is a very important predictor of sales and market success. Therefore, we will use product review rating as an outcome indicator to examine how the performance of a product can be affected by the type of ties and strength of ties. In the social network literature, the extent to which two nodes in a network share common interests, exchange messages and media content have constantly been adopted to establish a tie strength [2, 3, 12]. The more such data is exchanged between two nodes, the greater the relation and strength. Following these established measures, in the context of product reviews, we will use the extent of direct comparison between two products as an indicator of the strength of tie.

First of all, given that the strength of tie measures the extent of the competitive relationship between products, when a product is constantly being compared to a large number of competing products, it is an indicator that such a product is a popular candidate for consumers who focus on a given set of features in certain price ranges. Moreover, the stronger the overall strength of ties, the more likely that such a

product will stand out in the product comparison and competition. Hence, we predict that:

*H1: The review rating of a product is positively associated with the strength of ties.*

Secondly, based on Granovetter's "The Strength of Weak Ties" theory, weak ties may contain more useful information than strong ties as the source of information is more diverse [12, 14, 19]. For example, indirect relationships established within a social network can yield positive results, as demonstrated in the case of job searching. When applying this finding to the context of customer reviews, it is likely that the indirect comparison and the conclusions customers derive from weak tie relationships may offer additional insights on the quality or performance of the product being evaluated. For customers who want to compare two products involved in a weak tie relationship, they will find that the indirect comparison through a third product yields more valuable information that more than compensate for the lack of comparative reviews between the two weak tie products. Similar observations are echoed in the findings of the studies involving migrant social networks [21] and dating networks [20]. Hence, the comparative relationship derived from weak ties may play a more important role in influencing their perception on the overall quality or performance of a product. Therefore, we expect that:

*H2: The strength of weak ties has a stronger impact on product review rating than the strength of strong ties does.*

## 3. Research Methodology

The research methodology for identifying the strong and weak ties and computing the strength of ties consists of the following steps: collecting reviews, identifying reviews with comparison (using text mining), extracting comparative reviews and associated product features, establishing a tie through comparative reviews, and determining the weight of the feature for measuring the tie strength.

We have primarily focused on the reviews that explicitly compare two products. If there are at least one review that directly compares any two products, then a strong tie will be formed between them. Having more comparative reviews will increase the strength of tie (to be discussed later). If two products are not directly compared but both connect to a common third product through a strong tie relationship, then we define them as forming a weak tie. If there are no

comparisons of a product with another one among all the reviews posted for these two products, then there would be no connection between these two products in our network (absent tie). The comparisons could be based on any common features of both the products. All these steps are explained in the following subsections.

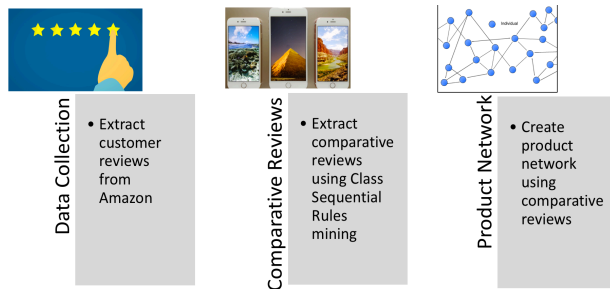


Figure 1a. Summary of Methodology

### 3.1. Data Collection

To ensure that we can derive a sufficient number of comparative reviews, we collected the user reviews of mobile phones of various brands in different price ranges from Amazon.com. As the largest e-retailer on the Internet, Amazon has a wide range of products in its comprehensive product categories and the most extensive customer reviews for these products. We choose to focus on mobile phones as this is one of the most popular and competitive product categories on Amazon. Due to close competition and the similarity in product specifications, consumers often provide comparison among similar products in this category, making it an ideal product category choice for our data collection. For each product in our dataset, we collected the product’s unique ASIN number (Amazon Standard Identification Number), review text, time of review, review rating, reviewer username, verified purchase flag, product name, product price, and various product specifications for products that appeared in the top-ranking lists during our data collection period between mid-April, 2018, through the first week of June, 2018.

We only focus on popular products as our main purpose is to apply our social network concepts to construct a product network characterized with ties. Since the non-mainstream products are less likely to emerge in comparative reviews, including these niche products will result in a product network with excessive absent-tie connections (no connections). Therefore, for ease of visualization and analysis, we

only focus on popular mainstream products. The dataset consists of 56,231 reviews from 137 products that belong to 32 brands. Out of these 56,231 reviews, we were able to extract 381 comparative reviews that have the competing product names clearly and completely mentioned in the review text. Table 1 provides a snapshot of the descriptive statistics of the collected customer reviews.

Table 1. Descriptive Statistics of Reviews Data

Key Variables	Count	Min	Max	Mean	Std. Dev.
Reviews	56,231	28	3,693	410	804
Price	137	\$44.5	\$1,143.6	\$305.5	\$246.50
Brands	32	-	-	-	-
Comparative reviews	381	0	44	3	2.74
Ratings	-	1	5	3.1	2.72
Average ratings range	137	2.12	4.3	3.1	2.72

### 3.2. Class Sequence Rule Mining

Analyzing the review text involves the choice of a suitable text-mining tool. We adopt the Class Sequence Rule (CSR) mining tool for our text analysis. CSR is part of the Association Rule Mining (ARM), a technique commonly used in text analytics. ARM is a rule-based mining, which specializes in classifying a set of inputs into the respective categories based on minimal set of rules. In CSR, the same technique is used, but in a more restricted manner. With text mining, rules can be built by associating set of words with the Part-of-Speech (POS) tags such as nouns (NN), verbs in base form (VB), verbs in past tense (VBD), and adjectives (AJ). For example, in the following sentence from a review, "when it comes to battery life, Samsung S9 is preferred over iPhone X". This sentence is a comparative sentence between two products Samsung S9 and iPhone X. When the stop-words like prepositions are removed and the rest of the words are POS tagged, the rule will turn the sentence to be "VB-NN-NN-NN-preferred-VBD-NN-NN". Each word with their POS tags will form an array that appears as: [(‘come’, ‘VB’), (‘battery’, ‘NN’), (‘life’, ‘NN’), (‘samsung’, ‘NN’), (‘s9’, ‘NN’), (‘preferred’, ‘VBD’), (‘iphone’, ‘NN’), (‘x’, ‘NN’)].<sup>1</sup> Preferred is retained in the sentence as it is one of the pre-selected comparative keywords around which the CSR rules are built. We considered all the common comparative words presented in [17] for extracting the comparative sentences and adapted their procedure for building a comprehensive rule set. Some examples of comparative keywords presented in [17] are: “more,

<sup>11</sup> A complete list is available at: [www.ling.upenn.edu/courses/Fall\\_2003/ling001/penn\\_treebank\\_pos.html](http://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html)

most, less, least, better, best, worse, worst, further/farther and furthest/farthest, beat, defeat, destroy, decimate, equal, equally, kill, lead, obliterate, outclass (out class and out-class), outdo, outperform, outplay, overtake (over-take), smack, subdue, subpar, surpass, top, trump, unmated, win, etc.”

To assess the effectiveness of the rules set and the comparative words list provided by [16], we created a training dataset out of all the text reviews and used this training dataset to improve the CSR algorithm. In the training process, we set the benchmarks by manually labeling the reviews with a ‘Y’ if there is a comparison of two products and with an ‘N’ if there is no such comparison. Next, we apply the Naïve Bayes algorithm to the training dataset to evaluate whether each sentence is a comparative sentence or not, based on the rules set. We then compare the results with the flags of ‘Y’ or ‘N’ that we have pre-assigned to each sentence. Initially, the rate of accuracy is less than 80%. However, after multiple iterations and training, the accuracy of the results reached 96%, which is a satisfactory accuracy rate compared to other similar studies [5, 18]. After verifying that the algorithm performed well on the training set, we then ran the algorithm on the rest of the records and marked them automatically with a ‘Y’ or ‘N’ labels.

### 3.3. Feature Extraction

Since a product may contain a variety of features on which two or multiple products can be compared, the product network built using the customer reviews is organized by different features. Building a product network by feature provides detailed insights mainly due to two reasons. First, not all product comparison reviews are based on all the features. Breaking the comparison down by features allows us to uncover any comparative relationship and maximize the number of ties that we can form within the network, thus revealing richer information from the network. Second, classifying ties based on features enable us to identify which feature contributes most to forming the ties. Product manufacturers can, thus, focus on improving the important features that their targeted consumers care most about.

The common features based on which the comparisons are made can be extracted using the CSR mining technique. Since the features of mobile devices can be described in a variety of ways, to make sure we capture these features correctly and merge similar descriptions for the same feature, we used the terms of features listed on the widely adopted electronic gadgets comparison website (gadgetsnow.com) [9], as used in [18] for extracting the features that are involved in the comparative texts. All the variants of

terms that are used in specifying a feature are included as query keywords when querying the review text. If the feature specific keywords are used in the review, then the review is marked to be related to that feature. Table 2 summarizes the feature keywords used in identifying the comparative reviews.

**Table 2. Features specific keywords**

Feature	Keywords
Screen	Screen, size, picture, quality, resolution, display, quad, glass.
Camera	Camera, pixel, rear, mp, hdr, portrait, autofocus, aperture.
Battery	Battery, charge, lithium, life, signal, mah, recharge, hour, day.
Software	Software, speed, fast, performance, ios, android, fingerprint.
Hardware	Hardware, sleek, edge, memory, expandable, storage, external, width, stylus, dimensions.

### 3.4. Building a Product Network through Strong and Weak Ties

The fourth step in our methodology is to build a product network using the product comparison reviews. We built the network on a two-dimensional graph. The vertical dimension is the price of the product and the horizontal dimension is the percentage of the total number of reviews of each feature (i.e. the number of reviews related to a given feature relative to the total number of reviews). The horizontal axis can also be considered a measure of the popularity of the product. Based on this graphical set up, the highest price product appears on the top of the graph, while the products with maximum percentage of reviews on a feature appear towards the far right on the graph.

In this stage, we identify the strong and weak ties based on the definitions provided earlier. For illustration, suppose that there are three products, P1, P2, and P3, in a given product category. There are direct comparisons between P1 and P2 as denoted by (P1, P2), and between P1 and P3 as denoted by (P1, P3), but there is no direct comparison between P2 and P3. In such cases, the connection between P1 and P2 and P1 and P3 are labeled strong ties, while the link between P2 and P3 is called a weak tie as they are both connected to P1 (see also Section 2.2).

As summarized in Section 2, the strength of a tie between two products A and B often depends on the extent of interaction. In the context of our study, we adapted this measure to capture how closely consumers view these two products and consider them as substitutes (competitors). Specifically, we proposed that the strength of tie ( $ST_{ij}$ ) for products  $i$  and  $j$  can be

calculated as the weighted number of times a product  $i$  and product  $j$  are being compared based on a given feature relative to the total number of comparisons that exist within a network, that is:

$$ST_{i,j} = \frac{\sum \text{Number of Common nodes of } i,j * \text{Feature Weight}_f}{\text{Total number of ties in the network}} \quad -- (1)$$

Since different features may have different impact on the comparison between two products, we will empirically determine the weight of each feature, and use them in the calculation of the strength of ties. To do so, we compute feature weight as a ratio of the total number of reviews related to this feature to the total number of reviews across all products in our dataset. As shown in Equation (2), feature weight essentially represents the importance of a given feature for all cellphones in our dataset.

$$\text{Feature Weight}_f = \frac{\text{Number of reviews on feature } f}{\text{Total Number of Reviews}} \quad --(2)$$

The combination of Equations (1) and (2) allows us to compute the strength of ties when the ties are based on a variety of features, which can be used to determine the extent of competition, or the profitability of a given market segment in the product space. When extending this concept to other contexts (i.e. traditional social networks characterized with varying forms of inter-personal communications and physical/virtual interactions), the features can be replaced by the types of interactions. For example, the strength of tie between two friends can be computed by the number of times these two persons communicate via phones, through Internet, or meet face-to-face. If someone is interested in learning about the strength of tie for a given type of communication or feature, then Equation (1) can be reduced to computing feature specific strength of the tie by removing the feature weight from the equation.

### 3.5. Regression Analysis

To test the hypotheses derived in the previous section, we will use a regression model to validate how the strength of strong and weak ties impact the review rating. The regression model is expressed below:

$$\begin{aligned} & \text{Product Review Rating} \\ & = \beta_0 + \beta_s(\text{Avg strength of Strong Ties}) \\ & + \beta_w(\text{Avg strength of Weak Ties}) \\ & + \text{control variables} + \epsilon \end{aligned}$$

In the above equation,  $\beta_0$ ,  $\beta_s$  and  $\beta_w$  are the coefficients of the intercept, the average strength of the surrounding strong ties, and the average strength of the surrounding weak ties, respectively, and  $\epsilon$  is the error term that captures unobserved factors that may contribute to the variation in the dependent variable. The control variables include the price of the product, number of reviews, brand, percentage of verified purchases, etc.

## 4. Results

We start by presenting the graphs of the product network generated using strong and weak ties. For ease of presentation, we break down the graphs by product features associated with the comparative review.

Figure 2 represents the network graph based on the comparative reviews about the products' screen. The ties in red are the weak ties, while the ties represented in black are the strong ties. The strengths of the ties are numerically expressed above the tie (i.e. line) in the network graph as network weights. Important observations and conclusions can be drawn by analyzing the network graphs in conjunction with the regression results. For example, i7 (iPhone X) from the screen features comparison graph of Figure 2 has more (strong and weak) ties when compared to r1 (Razer Phone). In other words, from this network graph, we observe that iPhone X is in a more centralized position or more commonly serve as the anchor point of comparison than the Razer phone. In fact, the product node i7 has more of both strong and weak ties in comparison to other product nodes; thus, we can conclude that i7 serves as the benchmark product in terms of the screen features, and is the strongest competitor for this given product feature. We can also immediately identify i7 (iPhone X) and p4 (Pixel) as direct competitors, and that i7 (iPhone X) and p3 (Pixel 2) and p5 (Pixel 2 XL), albeit there are newer products, are not being directly contrasted against iPhone X in terms of the screen features.

One observation from the graph is that strength of the weak tie helps in identifying a strategic position for launching a new product based on price range and features. For example, the strength of weak tie connection between Pixel 2 XL (p3) and Essential phone (ep1) is the highest, while the strong tie connection between these two products is currently absent (e.g. no customers have directly compared these two products). However, the high value of the strength of tie between these two products shows that there may exist some strong demand in the product space between these two products. Specifically, if a

product is launched with a price lesser than Pixel 2 XL and software features comparable to Essential Phone, then it will be a potential strong competitor. This is useful information for product manufacturers and designers.

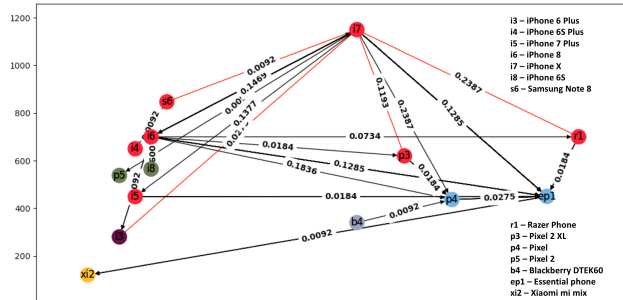


Figure 2. Product Network using customer reviews based on screen features

The product network graphs based on other product features are shown in Figures 3 through 6.

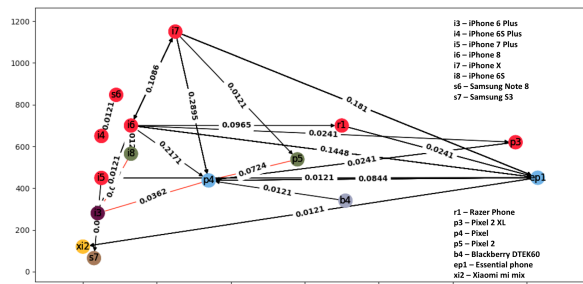


Figure 3. Product Network using customer reviews based on camera features

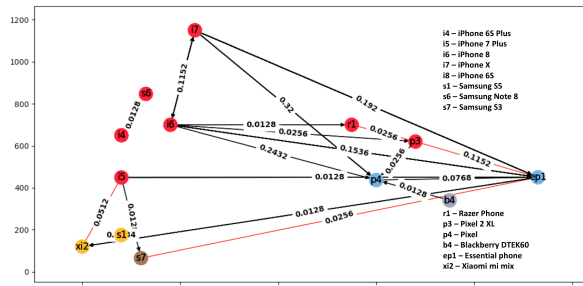


Figure 4. Product Network using customer reviews based on software features

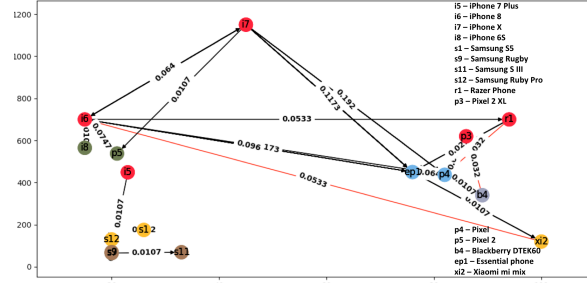


Figure 5. Product Network using customer reviews based on battery features

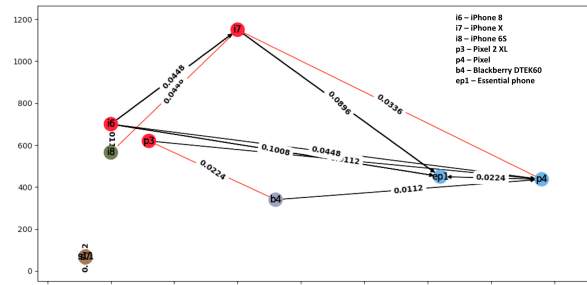


Figure 6. Product Network using customer reviews based on hardware features

In terms of the hypotheses testing, Table 3 shows the regression results using the average strength of the strong ties and the weak ties as the independent variables and the product review rating as the dependent variable. It can be seen from the table that the average strength of both the strong ties and the weak ties are positively associated with the review rating, and both coefficients are significant at  $p < 0.05\%$ ; thus, providing support for H1. Moreover, the overall contribution of weak ties is greater than that of the strong ties, which shows that the indirect comparison with the weak tie competitors actually has a greater influence on the product's review rating than that of strong ties. This confirms that the observation proposed by Granovetter's "The Strength of Weak Ties" theory also applies to the context of product social network. This finding implies that the search process can be greatly facilitated when they are trying to compare products that have not been contrasted directly, as the information derived from the weak ties can be more useful than if they are trying to compare the products directly.

Table 3. Regression Results



	Coefficient Estimate	Std. Error	Significance
Intercept	1.017	0.022	0.000**
Avg strength of strong ties	10.041	0.051	0.032*
Avg strength of weak ties	14.490	0.470	0.011*

\*  $p < 0.05$ , \*\*  $p < 0.01$

Table 4 shows the strength of all the weak ties created in the product networks. There are two significant weak ties (tie strength highlighted in blue) identified from the screen comparison reviews and two insignificant weak ties (where the tie strength is low, less than 0.05 in value). This table provides a comparative summary of the popular features that are being compared between products. Specifically, screen features are most widely discussed, followed by camera and software features. The weak ties formed due to the comparative reviews on hardware and battery features have the lowest values in tie strength, as they are not frequently mentioned in the reviews. This anecdotal evidence supports that screen and camera features play a crucial role in the overall product review. This has a practical implication as it helps product designers to align their efforts with customers' preferences while launching a new product.

**Table 4. Network graph results (Tie strength)**

(In)Significant Weak Tie	Source
(iPhone X, Razer Phone) – <b>0.2387</b>	Screen features
(iPhone X, Pixel 2 XL) – <b>0.1193</b>	
(iPhone X, Samsung Note 8) – <b>0.0092</b>	
(iPhone X, iPhone 6 Plus) – <b>0.0213</b>	
(Pixel, Pixel 2) – <b>0.0724</b>	Camera features
(Pixel, iPhone 6 Plus) – <b>0.0362</b>	
(Pixel, iPhone 6S) – <b>0.0362</b>	Software features
(Essential Phone, Pixel 2) – <b>0.1152</b>	
(iPhone 7 Plus, Xiaomi mi mix) – <b>0.0512</b>	
(Essential Phone, Samsung S3) – <b>0.0256</b>	
(Pixel 2 XL, Razer Phone) – <b>0.0256</b>	Hardware features
(Pixel 2 XL, DTEK60) – <b>0.0224</b>	
(iPhone 8, iPhone 6S) – <b>0.0448</b>	Battery features
(Pixel, iPhone X) – <b>0.0336</b>	
(iPhone 8, Xiaomi mi mix) – <b>0.0533</b>	
(iPhone 8, Pixel 2) – <b>0.0747</b>	
(Pixel, Razer Phone) – <b>0.032</b>	
(Pixel 2 XL, DTEK60) – <b>0.032</b>	

## 5. Conclusion and Limitations

The identification of strong and weak ties and their tie strength has important practical implications for a variety of stakeholders. For market operators, manufacturers, and sellers, locating strong ties allows them to visualize what products they are directly competing against, so that they can tailor their product design and marketing strategies to the characteristics of these direct competitors. The detection of weak ties also alerts them to potential competitors that are not clearly visible when using conventional indicators such as sales volume or review ratings. Moreover, if a product is surrounded by a large number of strong ties and does not stand out in the competition, it may need to consider repositioning its product in the product network to a different location where there are more weak ties or absent ties. In other words, in such a product network, areas surrounded by predominantly weak ties may suggest that this specific market segment is not fully saturated and has large potential for growth. Hence the distribution of weak ties can also be used to explore untapped market and expand business.

The effect of finding weak ties is particularly salient for consumers seeking to compare two distinct products but could not find comparative reviews for them. The weak ties will enable customers to identify indirect substitutes for their candidate products and find a common benchmark to compare such obscure substitutes through a common middle man (the common strong tie that they both connect with). This essentially reduces the search costs and improves the online shopping experience.

Our research provides a novel direction in which we can measure how a product will be compared if it is strategically launched in the market with a certain set of features and price. In this study, we identified that the screen, camera and software aspects play a crucial role in rating smartphones, which helps manufacturers learn about which of their product strength can be leveraged or improved. We were also able to visualize a unique first of its kind product network by utilizing social network theory “The Strength of Weak Ties”. We observe that the strength of tie is a significant predictor of a product’s review rating and this positive relationship applies to both strong and weak ties. Moreover, we find that the strength of weak ties has a more prominent influence than the strength of the strong ties, a finding that is consistent with the prediction in Granovetter’s “The Strength of Weak Ties” theory.

As with any empirical studies, our study is not without limitations, which also provide directions for future improvement and extension. As our primary focus of our study is to implement the concept of a

product social network and validate the strength of ties measure in such a product network, we limit the scope of data collection and only include popular products to prepare our input data for the network graph. We plan to enhance the overall model in our next study to extract more comparative reviews in an automated fashion which makes the study complete. Also, some reviews were dropped when the comparison does not provide the complete product name or involves brands instead of products (e.g: “My recent model of Moto Z is the best in its performance while compared to my previous moto version mobile”). Future studies may construct a nested model (products nested by brands) to investigate how the brand effect influence the outcome in such a product network. In addition, our sample data come from one leading electronic market, the results will be more generalizable if customer reviews can be collected from multiple online platforms and merged to create such a product social network. Lastly, in addition to strong and weak ties, other types of ties may exist in such product networks, and may have important influence on the outcome of product competition. Future extensions in this direction will contribute to a more thorough understanding of network ties in product competition.

## 6. References

- [1] Bapna, R., Gupta, A., Rice, S., & Sundararajan, A. (2017). Trust and the strength of ties in online social networks: An exploratory field experiment. *MIS Quarterly: Management Information Systems*, 41(1), 115-130.
- [2] Borgatti, S. P., & Feld, S. L. (1994). How to test the strength of weak ties theory. *Connections*, 17(1), 45-46.
- [3] Burt, R. S. (2002). The social capital of structural holes. *The new economic sociology: Developments in an emerging field*, 148, 190.
- [4] Castilho, D., Vaz de Melo, P. O. S., & Benevenuto, F. (2017). The strength of the work ties. *Information Sciences*, 375, 155-170. <https://doi.org/10.1016/j.ins.2016.09.009>
- [5] Chen, J., Huang, H., Tian, S., & Qu, Y. (2009). Feature selection for text classification with Naïve Bayes. *Expert Systems with Applications*, 36(3), 5432-5435.
- [6] Chen, K., Luo, P., & Wang, H. (2017). An influence framework on product word-of-mouth (WoM) measurement. *Information & Management*, 54(2), 228-240.
- [7] Choi, N., Kim, D., Goo, J., & Whitmore, A. (2008). Knowing is doing: An empirical validation of the relationship between managerial information security awareness and action. *Information Management & Computer Security*, 16(5), 484-501. <https://doi.org/10.1108/09685220810920558>
- [8] Church, E. M., Iyer, L. S., & Zhao, X. (2015). Using Interest Graphs to Predict Rich-Media Diffusion in Content-Based Online Social Networks. *Information Systems Management*, 32(3), 210-219.
- [9] Compare Mobile Phones - Latest Mobile Comparison by Price, Specification, Features, Performance & Reviews | Gadgets Now. (2018). Retrieved June 15, 2018, from <https://www.gadgetsnow.com/compare-mobile-phones>
- [10] De Bruyn, A., & Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing*, 25(3), 151-163. <https://doi.org/10.1016/j.ijresmar.2008.03.004>
- [11] Gee, L. K., Jones, J., & Burke, M. (2017). Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network. *Journal of Labor Economics*, 35(2), 485-518.
- [12] Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, 201. <https://doi.org/10.2307/202051>
- [13] Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives*, 19(1), 33-50. <https://doi.org/10.1257/0895330053147958>
- [14] Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 21.
- [15] Jain, R. (2017). Predicting Switching Intentions of the Indian Mobile Consumers. *International Journal*, 5(8).
- [16] Jindal, N., & Liu, B. (2006). Identifying comparative sentences in text documents (p. 244). ACM Press.
- [17] Kim, S., Kandampully, J., & Bilgihan, A. (2018). The influence of eWOM communications: An application of online social network framework. *Computers in Human Behavior*, 80, 243-254.
- [18] Kim, S. B., Han, K. S., Rim, H. C., & Myaeng, S. H. (2006). Some effective techniques for naive bayes text classification. *IEEE transactions on knowledge and data engineering*, 18(11), 1457-1466.
- [19] Lin, N. (1999). Building a Network Theory of Social Capital', 24.
- [20] Ortega, J., & Hergovich, P. (2017). The Strength of Absent Ties: Social Integration via Online Dating. *ArXiv:1709.10478 [Physics, q-Fin]*.
- [21] Ryan, L. (2011). Migrants' social networks and weak ties: accessing resources and constructing relationships post-migration. *The Sociological Review*, 59(4), 707-724.