Some Commercial Concerns of Amazon’s Community Forums
The Case of the Kindle

Allie Whitman
Marketing Department
The Pennsylvania State University
University Park, PA, USA

Bernard J. Jansen
Qatar Computing Research Institute
Hamad bin Khalifa University
Doha, Qatar
jjansen@acm.org

Abstract—Ecommerce consumers have many platforms for discussing products, one of which is the online product forum. These forums exist across the web on a variety of ecommerce shopping sites, including Amazon.com. However, the value of an online product forum for consumers, ecommerce, or the shopping site business is unclear. In this research, we examine 1,027 threads and 12,467 posts from the Amazon Kindle Forum, which is Amazon’s most active product forum. We use text analytics techniques to categorize the social media posts according to phases of buying decision process, simplified into pre-purchase and post-purchase activity. We use the metric of post helpfulness rating to measure the post value to the customer. Our analysis shows that customers primarily use the Kindle Forum for post-purchase activity. In addition, customers appear to find embedded links particularly helpful in the forum. Activity in the forum is related to the release dates of different Kindle models. Finally, results show that the forum can be used to identify customer complaints and product issues, so online discussion forums are a valuable tool for customer relationship management. The implications of the findings are that ecommerce companies could benefit from providing greater post-purchase customer support in their product forums.

Keywords — communities-of-practice; buying funnel; sales funnel; ecommerce shopping

I. INTRODUCTION

Of the online retailers in the US market, Amazon.com is one of the most successful at the time of the study. The company boasted billions of dollars in net sales, and the site has significant online consumer reach in the US. Although Amazon started out as a bookseller, it rapidly expanded to other product offerings, including its own line of e-readers and tablets, most notably the Kindle. Since the first Kindle was launched in November 2007, Amazon has released seven generations of the product, as well as the Kindle DX, Kindle Paperwhite, and Fire HDX (previously known as “Kindle Fire”). The Kindle holds a large share of the e-reader market.

Aside from selling products, Amazon.com also provides services for the customer community. One such service is the Amazon Discussion Forums. These forums are divided into two main categories: (1) community forums and (1) product forums. Community forums house general topics including science, politics, and music. Product forums contain discussions about specific product types.

For this research, we focused on the Kindle Forum, a product forum allowing customers to post comments related to Kindle products. Neither the Kindle Forum nor any other Amazon product forums have been previously analyzed, to our knowledge. However, we based our methodology on prior research done on other ecommerce forum sites [1-3], as well as previous Amazon seller forum studies [4, 5].

Our research objectives are (a) to determine the impact of certain forum elements, for instance, the relationship between types of posts and (b) the helpfulness of that post to a consumer, on ecommerce for both the consumer and the business. In doing so, we will analyze the effect of the Kindle Forum within a customer’s Buying Decision Process [6].

II. LITERATURE REVIEW

There have been several interesting news articles on the Amazon Discussion Forums, highlighting their uniqueness [7]. The primary question brought up by these articles is why does Amazon continue to allow irrelevant and often controversial content on its website? In addition, why do they put time and energy into moderating almost every individual discussion thread in order to actively restrain and/or delete comments? The answer is assumed that Amazon benefits greatly from the forums [3, 8]. However, the exact way it benefits, if at all, is unknown.

Relevant research pertains to studies of online communities, including forums, discussion boards, and brand communities—and how those communities affect ecommerce—as well as studies using data mining, text mining, and text analysis. We first examine at the concept of the online forum in general.

1 http://www.amazon.com/forum/amazon%20discussions%20feedback
Our research focuses on Amazon's Kindle Forum\(^2\), an online product forum for the Kindle and related products. An online product forum can be defined as a website where "rather than simply reading product reviews, a customer can post specific information he or she is looking for, and question the authors of specific data on particular issues" [9, p. 2]. Amazon's Kindle Forum meets this definition. Pitta and Fowler [9] developed a model of an online community's stages of development: potential, coalescing, active, dispersed, and memorable. In addition, they examined communities of interest, communities of practice, and communities of exchange. By their definition, the Amazon forums are a community-of-interest, incorporated into an underlying community-of-exchange, with lead users.

A community of interest is a group of people who share a common interest in a topic, such as the Amazon Kindle [9], which is an Amazon product. A community of exchange is one in which users provide information to one another, including asking and answering each other's questions or providing advice [9]. The Kindle Forum is nearly entirely user-based, with over 95% user-generated topics and posts. In addition to being a community of interest and exchange, the Kindle Forum is also an online brand community (OBC), characterized by discourse between users that allows virtual identities, discussion unconstrained by time or geography, and provides both functional benefits to users and firms [10, p. 227]. The idea of the OBC is the basis of our research goal to determine what functional benefits the Amazon Kindle Forum provides to consumers and Amazon (the firm).

A. Amazon Examined as a Data Source

While other aspects of Amazon have been studied before [5, 11-13], the Amazon product forums have not been extensively studied. Rooksby and Khajeh-Hosseini [5] examine Amazon's cloud computing service, Amazon Web Services (AWS), and they suggested how cloud computing providers can improve web forums to provide better customer support. Rooksby and Khajeh-Hosseini [5] refer to AWS forums as "socially organized, value laden venues for information" that "seem a convenient technology, both for provider and users" (p. 1). While the study looked at Amazon as a cloud computing provider, rather than an online retailer, its primary conclusions are still applicable to our research. For example, Rooksby and Khajeh-Hosseini [5] state that for a large-scale company like Amazon, direct contact with all users is unrealistic; instead, the company relies on secondary methods of support, namely the cheaper option of the forum. Forums "provide a familiar, many-to-many line of communication, and simultaneously build a temporal, public archive of problems and solutions" [5, p. 1]. The study also briefly examines hyperlinks, noting that they serve as either answers to a question or "serve not to provide extra details but as a strong indication that this problem is for the user to work out away from the forum" [5, p. 3].

Our research analyzes links, as well, but in a different way; we use the user helpfulness percentage of posts to measure the quantitative value of the links to the user.

Other studies look at Amazon's recommender system. Leino and Raiha [12] examine the importance of ratings and reviews in terms of related product recommendations, based on an earlier study showing ways recommenders influence consumer's opinions and buying behaviors [14], and developed a new algorithm for including ratings and reviews in the suggested products listing [12]. Linden, Smith, and York explained Amazon's recommender algorithm, as "Item-to-Item Collaborative Filtering" model [15, p. 1], but they did not create their own. Another study on the variance in product ratings determined that high variance in Amazon book ratings increase sales only if the average rating is low. A similar study examines the role of the product description on the purchasing decision for electronic devices for people ages 13 to 35 (Le Marechal, 2015). This study addresses another form of the Buying Decision Process [6], called the Circular Decision Making Process. The results of the research showed that people go online to compare price and learn about a product, as well as to complete the final sale [16].

B. Textual Analysis Research in Online Forums

Other studies have analyzed the results of text mining in online customer reviews. At least three prior studies use text mining to analyze non-Amazon customer reviews [8, 11, 13]. Dellarocas examine secondary data to see the impact of a company manipulating reviews and determined that the effect is based on a firm's true quality and a consumer's original intent to buy [11]. Moe and Trusov [17] look at a home goods website's online reviews and were able to classify consumer's buying behavior by their online review interactions [8]. Zhu and fellow researchers look at GameSpot reviews and found that consumers' behavior differed based on experience level and product popularity [18]. Still more researchers specifically examine Amazon's customer reviews [12, 19]. These researchers primarily develop algorithms for text mining customer reviews; since our purposes are not to create new algorithm but to do a direct analysis, such research is relevant in methodology than in wider reaching managerial application.

Klaus [20] evaluate Amazon's online customer service experience, determining that the two main elements dictating whether someone will have a positive or negative experience are functionality and psychological factors [20]. This was similar to two other research studies. The first study shows the distribution of positive and negative Amazon product reviews for children's tutus over time [21]. The result was that positive reviews tend to be evenly dispersed, but negative reviews tend to occur at later steps in the review process. The second study looks at the factors contributing to Amazon product reviews level of helpfulness. Researchers divide the helpfulness down into levels of credibility, levels of diagnosticity, and product type: search or experience [22]. Other factors that might affect the customer's experience are covered in a study that examines whether disclosing whether a reviewer has actually

\(^{2}\) http://www.amazon.com/forum/kindle
purchased the product through a "Verified Purchase" stamp has a significant effect on helpfulness ratings of reviews, separated into "experience" and "search" reviews [23]. The findings show that the Verified Purchase badge does correlate with higher helpfulness ratings, but the researchers did not look into its possible effect on product sales [23].

C. Gaps in Prior Research

There is a notable lack of research done specifically on Amazon’s website marketing, one reason most likely being that Amazon makes very little sales data available publicly. It is therefore difficult to measure quantitatively the effectiveness of Amazon’s marketing tactics. Other Amazon studies look only at Amazon’s general business practices, supply chain innovations, and IT structure; these are not relevant to our research. Several studies have examined online forums as social communities.

Given this prior research, there are several open questions. Why do customers visit a product forum in the first place? How can a company optimize a product forum for both customers and themselves? These are some of the questions that motivate our research.

III. RESEARCH QUESTIONS

For this research, we focus on one primary research question.

RQ1: What value does the Kindle Forum provide to customers, including those who have not posted in the forum but have read posts? Specifically, by value, does the Kindle Forum provide primarily:

A) Pre-purchase services (product information, product specifications, price comparisons, etc.) to aid the consumer in making a purchase, and/or

B) Post-purchase services (customer feedback, product questions or help, complaints, product suggestions, etc.)

For this research questions, we examine metrics such as the percentage of pre-purchase vs. post-purchase posts, and the percentage of different types of posts compared by the helpfulness ratings, in order to address the questions.

IV. METHODOLOGY

We first provide an overview of Amazon.com and its discussion forums, including the layout and function of the forums, discussion threads, and user community. We define our terminology in Table 1.

A. Data Collection Site

Any Internet user can view the Amazon discussion forums, but only registered Amazon customers can participate (e.g., post) in the forums. The majority of the forums were created by Amazon itself, such as the Politics, Music, Books, Religion, and Kindle forums.

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>TERMINOLOGY AND KEY CONCEPTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>Forum</td>
<td>Collection of discussion threads grouped under a single category</td>
</tr>
<tr>
<td>Discussion thread or thread</td>
<td>Collection of comments or posts and their replies grouped under a single topic</td>
</tr>
<tr>
<td>Comment or post</td>
<td>Text comments including the comment box (username, timestamp, adds to discussion, permalink)</td>
</tr>
<tr>
<td>Original post</td>
<td>An independent comment or post that was not created as a nested reply to another post</td>
</tr>
<tr>
<td>Reply</td>
<td>A nested post created as a reply to another post</td>
</tr>
<tr>
<td>Link</td>
<td>A hyperlink found within the content of a post</td>
</tr>
<tr>
<td>Embedded link or Amazon link</td>
<td>An Amazon.com product page hyperlink that is displayed underlined and in blue font</td>
</tr>
<tr>
<td>Helpfulness rating</td>
<td>The percentage indicating how helpful other users found the post</td>
</tr>
</tbody>
</table>

Within these forums, both Amazon representatives ("Amazon Official" badge holders, or "Amazon Community Team") and customers can begin their own discussion threads. Amazon-made discussion threads are called "Announcements," and threads can either allow or disallow comments. Announcements always appear at the top of the list of discussion threads within a forum. The majority of discussion threads, however, are user-generated.

Fig. 1. Layout of the Amazon community forum with thread name, announcements, user posts, and aggregate statistics.

1) Anatomy of a Discussion Thread

Figure 1 depicts a screenshot of the Kindle Forum. Each page of the forum includes 25 discussions. For each discussion, the title of the discussion, number of posts, and time of latest post are shown. The sidebar also displays the total number of discussions contained in the forum. The forum includes official Amazon posts called Announcements. These are discussion threads started by Amazon employees and moderators called “Amazon Officials.” Announcements typically address new product offerings, updates, or popular customer complaints. Similar to customer-created discussion threads and posts,
Announcements allow users to reply to the post, as well as submit a helpfulness rating.

1) Data Collection

In total, the data scraped comprised Kindle forum threads, posts, and customer profiles from 2009 to 2015. Data was mined using the web crawling platform, Import.io. First, an API was made to collect the links on each page of discussion titles. The title links were then exported to a customized crawler that scraped the posts from each discussion thread. The information scraped includes: permalink URL, post date, post text, links contained in the text, contributor name and profile link, reply vs. original comment post label, report abuse rating (helpfulness rating). The total number of posts crawled was 12,467. A total of 4,292 unique participants posted in the forum, and each user posted an average of about 3 posts. The highest number of posts made by a single user was 447, and the lowest was 1. One post was the most frequently occurring number of posts made.

- Total number of discussion threads: 1,027
- Total number of posts: 12,467
- Total number of announcements: 4
- Total number of unique participants: 4,292

Once scraped, we analyzed the data by first doing term level analysis to determine the highest frequency words and phrases. This information was used to group discussion posts into several categories based on post topic, as will be described later. Additional labels, such as "helpful" (rating of 50% or higher) or "unhelpful" (rating of 49% or lower) were applied to all comments. We also categorized the posts as either Pre-Purchase or Post-Purchase as based on the consumer behavior Buying Decision Process [6] using the methodology described below. Pre-Purchase refers to any activity prior to purchasing a Kindle product, while Post-Purchase denotes activity after buying a Kindle product.

Consumer behavior theory states that there are five primary steps in the Buying Decision Process [6].

- The first step is “Need Recognition,” also known as “Problem Recognition,” during which the consumer realizes he has a problem that must be solved or a need that must be met. This occurs when there is a discrepancy between the actual state and the desired state.
- Next is “Information Search,” during which the consumer researches information on how he can meet his need or solve his problem. “Need Recognition” can be caused by internal or external stimuli.
- The consumer then completes the “Information Search,” to look for information about possible solutions. Like “Need Recognition,” information can come from internal or external sources.
- The consumer then completes an “Evaluation of Alternatives,” to compare possible options based on attributes (feature, functionality) and perceived value.
- Next, the consumer completes the “Purchase Decision” step by deciding whether to buy or not to buy the product.

- Finally, “Post-Purchase Behavior” refers to the consumer’s evaluation of the purchase as well as the consumer’s level of satisfaction or dissatisfaction. Returning a product to receive a refund or asking for customer support would both fall under “Post-Purchase Behavior.”

Below are examples of the Buying Decision Process applied to the Kindle forum discussions:

- Need Recognition: The consumer needs a portable way to access and read many books.
- Information Search: The consumer asks friends for advice and discovers e-readers. The consumer does internet research on the various types of e-readers that exist, including the Kindle Fire and the Barnes and Noble Nook.
- Evaluation of Alternatives: The consumer compares the Kindle Fire to the Nook and decides the Kindle Fire better fits his needs.
- Purchase Decision: The consumer decides to purchase the Kindle Fire.
- Post-Purchase Behavior: The consumer is satisfied with his purchase, but needs help installing apps on his Kindle Fire, so he contacts Amazon Customer Support.

We categorize discussions in the Kindle Forum based upon the step in the Buying Decision Process with which they most corresponded. For the purposes of our research, we simplify the Buying Decision Process into two groups.

Take a discussion titled “Kindle Fire vs. Kindle Paperwhite,” for example. This thread asks other customers to discuss the merits of each product, but it could fall under either “Information Search” or “Evaluation of Alternatives,” since the customer is simultaneously comparing two products and searching for information on those products.

To facilitate analysis, we group the first three steps into Pre-Purchase and the last step into Post-Purchase. Post-purchase behavior might include any of the following: customer complaints, product reviews, suggestions directed at Amazon, or questions about how to use or fix the product. The actual Purchase Decision step is not illustrated in the forum.

V. RESULTS AND ANALYSIS

To perform our analysis, each discussion thread was first filtered and tagged based on keywords. Keyword tags are based on text analysis concerning the frequency of words and phrases. We assume that the posts contained under each thread relate to the original topic, and thus apply conclusions from the thread title analysis to the entire forum.

We compiled the most frequently occurring phrases in the thread titles, and identified three notable categories: problem, app, and model name. Threads falling under the problem category include complaints about functions of the Kindle products, or questions about the Kindle products. Threads falling under the app category include mentions of any
application not including a physical Kindle product. Thread titles mentioning the specific name of any Kindle models (not including the generic “Kindle”) were grouped into the model name category. Table II provides definitions and examples.

**TABLE II.**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>Complaints Description</td>
<td>Keyword Tag Examples</td>
</tr>
<tr>
<td>Problem</td>
<td>about Kindle problems, issues or malfunctions that need to be resolved</td>
<td>won’t turn/will not, can’t/cannot, how to, why, issue/issues, problem/problems, how can, missing, not working, broken, unable, help, questions, no longer, not, error, fail, suggestion</td>
</tr>
<tr>
<td>App</td>
<td>Applications not including e-readers</td>
<td>iPad app, phone app, Android app, desktop app, ereaderiq, Kindle for PC, Kindle for Mac</td>
</tr>
</tbody>
</table>

**TABLE III.**

<table>
<thead>
<tr>
<th># of Words</th>
<th>Most Frequently Occurring Phrases</th>
<th>% of Total</th>
<th>Occurrences</th>
<th>Category Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>kindle fire</td>
<td>15.8%</td>
<td>227</td>
<td>model name</td>
</tr>
<tr>
<td>2</td>
<td>fire hd</td>
<td>4.9%</td>
<td>70</td>
<td>model name</td>
</tr>
<tr>
<td>2</td>
<td>my kindle</td>
<td>4.6%</td>
<td>66</td>
<td>none</td>
</tr>
<tr>
<td>3</td>
<td>kindle fire hd</td>
<td>4.3%</td>
<td>62</td>
<td>model name</td>
</tr>
<tr>
<td>2</td>
<td>on kindle</td>
<td>4.0%</td>
<td>58</td>
<td>none</td>
</tr>
</tbody>
</table>

To answer our first research question— *what is the value of the Kindle Forum to Amazon customers?* — we analyzed the forum thread titles based on whether they are post-purchase or pre-purchase. We categorized every discussion thread by tagging it either pre-purchase or post-purchase, by examining keywords and the actual post text to determine context (see Table III). All discussion thread titles containing Problem-tagged phrases were categorized into post-purchase. Next, remaining threads were categorized by manually evaluating each title and grouping all threads where the user is specifically asking about recommendations for purchasing a new Kindle model or deciding between different products, into pre-purchase. The last remaining threads were categorized by reading through text and comments to ascertain whether or not the user bought a product or is looking to buy a product.

Our results reveal that the majority (87%) of the discussions were post-purchase, while the remaining 13% were pre-purchase. Of the post-purchase discussions, one large reason for customers visiting the Kindle forum appeared to be to resolve problems. Of the posts, 26% of the discussions were related to complaints, issues, or malfunctions with Kindle products. Some examples include:

- **Has anyone figured out a way to access your photos uploaded to cloud on the Fire?**

- **Same problem, let me put in wifi info, said to plug to charger, did that, but never got anything after that even after reboot. Black screen with Kindle Fire across it. Took back got replacement and same problem, maybe a bad batch??? Am getting my money back today, twice nonfunctional is enough for me.**

While we can conclude that users primarily use the Kindle Forum for post-purchase discussion, there was no significant difference between the helpfulness ratings of post-purchase posts and pre-purchase posts.

We also graphed the number of mentions of nine Kindle Models (Kindle 2, Kindle Fire, Kindle DX, Kindle Keyboard, Kindle 4, Kindle Touch, Kindle Paperwhite, Kindle Voyage, and Kindle HDX) from year to year. We see that the first mentions of all models, except Kindle 4 and Kindle HDX, occurred after the release year. Kindle Fire had the highest number of mentions in a given year. Therefore, this analysis supports our findings that the majority of discussion in the forums was post-purchase.

In addition, links proved to be helpful for users, with 5% of posts (591 posts) containing at least one link. Of those posts containing links, 46% contained embedded product links. This suggests that embedded product links provide some value to the user. There were also no non-Amazon links (links that led to external sites). All links were Amazon domain, although some were not product links but instead links to other Amazon pages, such as help pages, customer service pages, or other forum pages. Analysis showed that posts with links were generally more helpful than posts without links. On average, posts that contained links were rated 12% more helpful than posts without links. There was a significant effect for links, t (646) = 6.53, p = 0.000, with posts with links receiving higher helpfulness votes than posts without links. In addition, the most frequent helpfulness rating of posts with links was 100%, while the most frequent helpfulness rating of posts without links was 0%. However, if a post had at least 1 link, there were there was no correlation between the number of links contained in a post and the post’s helpfulness score, nor between the number of links contained in the post and the post’s publish date.

**VI. DISCUSSION**

The primary value of the Kindle Forum to Amazon customers is for post-purchase services. Upon analysis of the Kindle Forum data, it seems that the majority (70%) of forum threads and posts were related to post-purchase topics, while only 13% were related to pre-purchase topics. This suggests that the primary value of the Kindle Forum for customers is as a post-purchase tool. We also analyzed the relationship between the release dates of various Kindle products, and the frequency of those products’ mentions in the forum, finding that the bulk of mentions of a certain model fell the year of, or even years after their release dates. This reaffirms the fact that most discussion in the forum was of post-purchase nature.
Links also proved important to the customer in terms of helpfulness of posts.

Like all research, there are strengths and limitations. The primary limitations of research were the lack of demographic data from users, the absence of purchase validation (such as the "Verified Purchase" stamp on reviews) in the forum, and the rudimentary nature of the text analysis used. With more sophisticated text analysis techniques, we could determine the most common product problems in more detail. Strengths of this research were the sample size (roughly 14,000 comments) and the fact that very little prior research (if any), exists on the Amazon Kindle or other product forums. The Kindle Forum is particularly unique, because it represents one of Amazon's proprietary and core products. To our knowledge, no other studies have looked at an e-retailer's own in-house branded product forum or review section.

The results of our research have several managerial implications. Amazon could better utilize the Kindle Forum by providing sub-forums for each Kindle Model to help users find help faster. Amazon could also use the Kindle Forum to identify problems with the product to both respond to customer complaints and improve upon current models in the future.

VII. CONCLUSION AND FUTURE RESEARCH

Our research determined that the Amazon Kindle Forum primarily provides value to customers in the form of post-purchase services, while providing value to the company in the form of customer service and product development. In terms of future research, we will examine the relationship of the Amazon Kindle Forum and Kindle product reviews around topic [24]. For example, we could determine whether the product reviews are where customers go for pre-purchase information instead of the forum. We could also compare the Kindle Forum to other product forums to determine whether our conclusions are replicable.

REFERENCES


