Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2009 Proceedings

Americas Conference on Information Systems (AMCIS)

2009

Automatically Learning User Needs from Online Reviews for New Product Design

Thomas Y. Lee
University of Pennsylvania, thomasyl@wharton.upenn.edu

Automatically Learning User Needs from Online Reviews for New Product Design

Thomas Y. Lee
Operations and Information Management Department
University of Pennsylvania, The Wharton School
thomasyl@wharton.upenn.edu

ABSTRACT

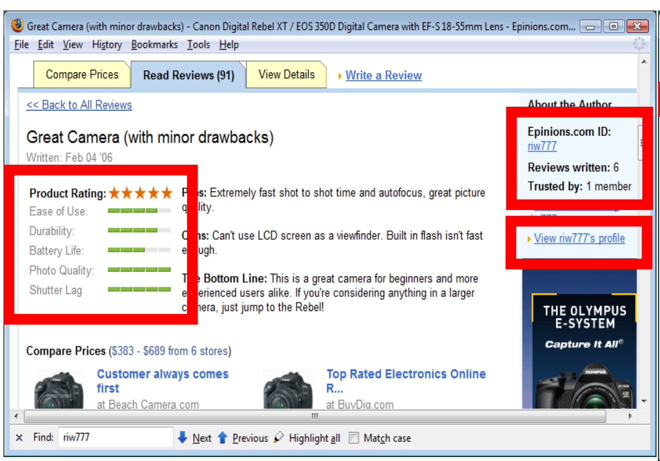
The traditional product design process begins with the identification of user needs (Ulrich and Eppinger 2008). Traditional methods for needs identification include focus groups, surveys, interviews, and anthropological studies. In this paper, we propose to augment traditional methods for identifying user needs by automatically analyzing user-generated online product reviews. Specifically, we present a supervised, machine learning approach for sentential-level adaptive text extraction and mining. Based upon a set of 9700+ digital camera product reviews gathered in January 2008, we evaluate the approach in three ways. First, we report precision and recall using n-fold cross-validation on labeled data. Second, we compare the recall of automated learning with respect to traditional measures for identifying users and their respective needs. Third, we use multi-dimensional scaling (MDS) to visualize the competitive landscape by mapping existing products in terms of the user needs that they address.

Keywords

Information extraction, new product development, supervised machine learning, product reviews

INTRODUCTION

In their reference work on Product Design and Development, Ulrich and Eppinger note that 80 to 90% of successful market innovations follow a traditional user-pull design process (Ulrich and Eppinger 2008). That process begins with customer needs. The traditional approach to discovering user needs involves focus groups, surveys, one-on-one interviews, or even anthropological, observational studies (Griffin and Hauser 1993). However, customers today willingly volunteer their thoughts and opinions online. Much of the current research involving customer reviews centers on structured, quantitative variables. Examples include price, the numerical rating of a product and/or its performance, and the authority of the reviewer as established by reviewer rating or number of past reviews (see Figure 1).



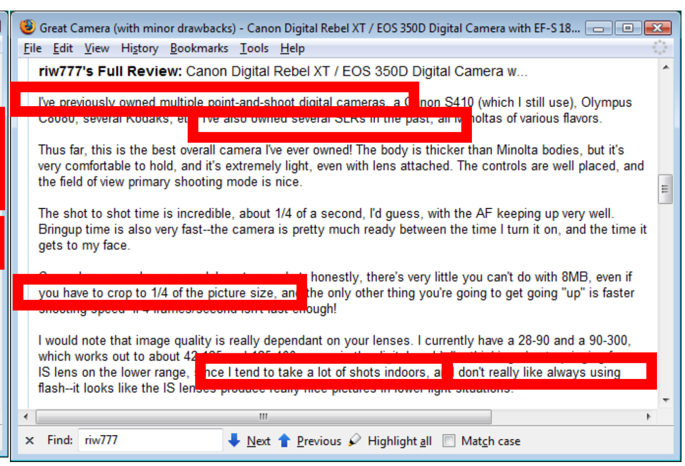


Figure 1. Traditional Review Analysis

Figure 2. User needs as articulated in reviews

In addition to these structured values, however, users often share a tremendous amount of additional knowledge in the text of their reviews. In particular, users include *behavioralistic* variables (Urban and Hauser 1993) describing how they actually use the product. Users also report *psychographic* (Urban and Hauser 1993) variables (activities and interests) that define the context in which the product is used. In short, users describe their *needs*. For example, the customer might reveal that they own, or have owned in the past, several cameras. The customer might articulate how they use the product: they crop photos, they take pictures indoors, and they do not like using the flash (see Figure 2).

We propose to augment traditional methods for learning user needs by automatically processing online product reviews. Why analyze the online review space? Automated methods complement existing methods in at least two ways. First, automation decreases the cost of acquiring needs while increasing the breadth of needs elicited. Traditional needs assessment involves the careful selection of a sufficient quantity of representative consumers in different market segments to ensure that all needs are accounted for. Griffin and Hauser (1993) mapped the decreasing returns from interviewing additional consumers for eliciting user needs. Automation can increase the comprehensiveness of traditional needs assessment at the minimal cost of simply processing additional reviews and/or processing reviews from different sources (representing different consumer groups). Second, automation facilitates analysis over time. Even if automated methods do not identify needs already discovered by traditional methods, automated methods are easily and inexpensively repeated. Tracking changes in user needs over time using traditional methods is laborious. For rapidly evolving products and services, automated analysis of product reviews promise a faster, simpler approach for remaining abreast of changes in the consumer marketplace.

We present a supervised, machine learning approach for sentential-level adaptive text extraction and mining. After first reviewing the related literature, we detail our proposed process for automatically learning needs from reviews. Next, we discuss the results of several preliminary evaluations based upon a set of several thousand digital camera reviews from Epinions.com. The paper concludes with limitations and future work.

RELATED WORK

This research integrates several bodies of prior work including research on needs-based analysis, sentiment mining, product feature extraction, and adaptive text extraction. The prior work on needs-based analysis encompasses work both in marketing and product design. In their Trusted Advisor project, Urban et al. develop an on-line truck buying service that recommends products based upon a profile of prospective needs (e.g. how many passengers will you carry, how much cargo do you have to haul) and invites users to "meet other people like me" through an online forum (Shankar et al. 2002). However, the knowledge base by which users are matched to products based upon needs is constructed manually. In this work, we discuss automated needs identification for the purpose of automatically generating such a knowledgebase by automatically processing online product reviews. The Trusted Advisor is one instance of a Customer Decision Support System as envisoned by Orman (Orman 2007). While Orman discusses the potential for using ontology-based strategies that might be used to construct and reason over a knowledgebase of needs, we focus on adaptive text extraction and implement a prototype system to automatically extract needs. In the context of marketing, a needs-based analysis may also be framed as an extension to traditional recommender systems. Needs such as "date-night" or "family outing" might serve as additional dimensions in a multi-dimensional recommender system (Adomavicius et al. 2005). However, the work on recommender systems rely upon user input in the form of transaction records to acquire the data and assumes prior knowledge of which dimensions (e.g. which needs) are important to query the user for. In this work, we acquire the data automatically from product reviews and identify "important" needs based upon what appears in the product reviews.

Paralleling the work needs-based analysis in marketing is the work on needs-based analysis and product design. Urban and Hauser extend their work on the Trusted Advisor with a second knowledgebase that informs a Virtual Engineer (Urban and Hauser 2003). The goal of the virtual engineer is to query users in order to learn about needs that are not met by existing products. The Virtual Engineer again assumes an existing set of needs, defines products in terms of a vector of attributes, and logs user interactions with the virtual engineer to assess demand for new products that lie at the intersecton of needs and product attributes where no product currently exists. Similarly, Randall et al. discuss needs-based design in the context of mass-customization (Randall et al. 2007). Given a pre-defined set of needs, a pre-specified multi-attribute utility function maps need preferences to product attributes. By entering needs-preferences, users design customized laptop configurations.

While the prior work on needs-based analysis assumes the manual acquisition of data, the textmining and economic communities have developed automated methods for analyzing online product reviews. Early work looked at the accuracy of predicting a numerical rating (of movies, consumer services, or automobiles) based upon the sentiment terms used in the text of the corresponding reviews (Turney 2002). The economic community applied Turney's approach to measure the economic impact of reviews based upon their positive and negative sentiment (as opposed to numerical ratings) (Das and Chen 2007; Pavlou and Dimoka 2006). In addition to sentiment, text features such as review length (Chevalier and Mayzlin 2006) or readability and polarity (Ghose et al. 2006) is shown to improve the resolution of such economic analyses. Our work complements current models by proposing a complementary set of new variables: user needs.

In contrast to our focus on needs, computer science researchers observed that sentiment terms tend to modify product attributes (e.g. love the zoom on this camera) and leveraged a vocabulary of sentiment terms with term frequency analysis to automatically identify product attributes (Hu and Liu 2004; Nasukawa and Yi 2003). By also including Turney's use of Pointwise Mutual Information (PMI), researchers were able to improve the accuracy with which they could identify product attributes (Popescu and Etzioni 2005; Scaffidi et al. 2007). Given automatically identified attributes, researchers used sentiment terms to automatically rate a product's performance on specific attributes (Liu et al. 2005) and extended that work to automatically measure the economic price-premium of positive and negative comments (Archak et al. 2007) Existing methods for identifying attributes, however, are frequency-based and do not distinguish between attributes and needs. For example, state of the art results in attribute identification (Scaffidi et al. 2007) processed consumer reviews of barbecue grills and found that one prominent attribute of grills is "hamburgers." In this work, we complement research in the computer science community by articulating a distinct problem: identifying needs. We differentiate between product attributes and user needs by proposing a method that explicitly addresses needs identification.

Our focus on the automated identification of user needs is unique. However, our technical approach builds upon the literature in adaptive text extraction. Our basic approach to discovering needs in user reviews follows the intuition outlined in WEIN (Kushmerick 1997). Based upon a hand labeled set of training examples, we learn patterns of text that commonly precede or follow a need. RAPIER (Califf et al. 1999) models the process of discovering such patterns as a greedy, hill-climbing, beam-search over the space of all training examples. WHISK (Soderland 1999) models the same search space, but uses a different objective function: minimize the Laplacian estimate of the standard extraction error. STALKER (Muslea et al. 2001) introduces the intuition that the location of "where" an item appears within the text (e.g. the introduction or conclusion) provides additional context for defining patterns. We define a new process for defining common patterns that is inspired by a combination of these different approaches. Although we also attempt to minimize the Laplacian, we define our search in evolutionary terms rather than hill-climbing. We also use location landmarks but use linguistic syntax (e.g. sentence subject, sentence predicate) rather than semantic sections (e.g. introduction, conclusion).

APPROACH

We describe a process for learning patterns to extract user needs from online product reviews based upon a supervised machine learning algorithm. Beginning with the raw text of a review, we first apply some preliminary linguistic preprocessing. A set of extraction patterns is initialized, trained, and tested based upon the preprocessed data. The resulting extraction patterns are applied to new reviews.

Linguistic Preprocessing

Linguistic preprocessing involves four steps that are highlighted in Table 1. Step 1. Beginning with raw text, words are labeled with their grammatical part-of-speech (POS). Step 2. The tagged sentences are then decomposed into subject-verb-object (SVO) triples. A single sentence may include several SVO triples as in compound sentences, prepositional phrases, etc. Step 3. Words are then lemmatized to their root forms to account for plural forms, past v. present tense, etc. Step 4. Finally, for purposes of learning user needs, we define an s-filter (subject-filter). SVO triples are filtered based upon the contents of their subject. Specifically, only those sentences with a personal pronoun (e.g. I, he, she, our, etc.) or the proper name of a product or brand name (e.g. Canon, Nikon, EOS, etc.) is used in training. Put differently, as a location landmark used in training [Muslea], we only look for user needs in a sentence predicate and prefilter sentence predicates based upon their corresponding sentence subject. We report sensitivity results to validate this heuristic as part of our evaluation.

Table 1. Linguistic Preprocessing

Step	Task	Example
	raw text	… since I tend to take a lot of shots indoors ...
1	POS Tag	… I(PRP) tend(VBP) to(TO) take(VB) a(DT) lot(NN) of(IN) shots(NNS) indoors(NN)
2	chunk	[S: I] [V: tend to take] [O: lot] [O: of shots indoors]
3	lemmatise	I tend to take lot of shots(shot) indoors(indoor)
4	s filter	I (PRP)

Training and Testing

For training and testing purposes, we begin with a set of randomly selected SVO phrases. The needs in each SVO phrase of this training/testing set are manually labeled: Each phrase is separated into a prefix, a *need*, and a suffix. The prefix and suffix serve as natural language text delimiters for the need (See Table 2). Note that it is possible for the suffix to simply be the end-of-phrase marker.

Table 2. Labeling SVO phrases

	Prefix	Need	Suffix
Parent 1	I tend to	take a lot of shots indoors	
Parent 2	I can	take several shots in succession	

After splitting labeled data into training and hold-out sets, training and testing begins by initializing a population of candidate extraction patterns. A pattern is initialized by transforming a labeled phrase into a three-dimensional regular expression. The three dimensions correspond to the raw text, the lemmatized form of the text, and the POS tags for each word. The three dimensional regular expression extraction patterns for the example SVO phrases in Table 2 are depicted in Table 3. Notice that the phrases are now transposed to better display each of the three dimensions in the prefix, *need*, and suffix. In regular expression terms, notice that the end-of-phrase marker is translated into a non-greedy match to the end-of-sentence. The *need* is captured in the regular expression as a back-reference. Intuitively, a need is extracted if any one dimension of the prefix pattern, one dimension of the *need* pattern, and one dimension of the suffix pattern all match.

Table 3. Initializing, Evolving, and Applying a Population of Patterns

		Parent 1	Parent 2	LCS	Extract
Prefix	Txt	i tend to	i can	i .*?	you can
	Lem	i tend to	i can	i .*?	you can
	POS	PRP VBP TO	PRP MD	PRP .*?	PRP MD
Need	Txt	take lot of shots indoors	take several shots in succession	take .*? shots .*?	take sharp macro images handheld with flash
	Lem	take lot of shot indoor	take several shot in succession	take .*? shot .*?	take sharp macro image handheld with flash
	POS	VB NN IN NNS NN	VB JJ NNS IN NN	VB .*? .*? NNS .*? NN	VB JJ NN NNS JJ IN NN
Suffix	Txt	.*?	.*?	.*?	thanks to cameara
	Lem	.*?	.*?	.*?	thank to cameara
	POS	.*?	.*?	.*?	NNS TO NN

From an initial population of patterns, we evolve the population by randomly selecting any two parents and crossing those patterns to generate a child pattern. The goal of evolution is to produce a child that is more general than either parent and can replace both parents in the population. The cross-over process between two parents is applied independently to each part of the pattern (prefix, *need*, suffix) and to each dimension of each part. Specifically, we look for the longest common substring (LCS) between the two parents, checking each part and each dimension separately. Common substrings are separated by non-greedy wildcard matches. The process is illustrated in the third column of Table 3. Because any two patterns may have more than one LCS, we currently generate candidates from only the first such LCS. This is similar to the beam search employed in RAPIER with a beam width of one (Califf et al. 1999). Extending this to random combinations of alternate LCS would be straightforward in an evolutionary framework.

To determine whether a child can replace both parents, we score each parent and each child using the Laplacian estimate of the standard error. Testing each pattern against the labeled training population, let c be the number of correctly matched need phrases and e be the number of errors. The Laplacian is then defined by:

$$Laplacian = \frac{e+1}{c+e+1}$$

A child replaces both parents as a function of both the Laplacian and coverage. Our current implementation uses a very conservative decision rule. A child replaces a parent only if the child scores strictly lower than the parent and the child covers at least as many labeled test phrases as the parent. Coverage refers to the number of labeled SVO phrases matched (correctly or incorrectly) by a pattern. Intuitively, each pattern in the initial population begins with a Laplacian of 0.5 (alternately, a precision of 1.0 where precision is defined as $c/(c+e)$). Using our conservative decision rule, the precision of any surviving child pattern remains at 1.0 at the possible penalty of compromising recall. Where p represents a pattern in the population, the overall objective of evolution is defined by the function:

$$min \sum_{\forall p} \frac{e_p + 1}{c_p + e_p + 1}$$

Evolution halts either after a maximum number of iterations or when the change in the objective function fails to exceed some threshold over a fixed number of iterations.

Once the final population of patterns is established, we apply the patterns to a new set of reviews. For example, in Table 3 column 4, we see how the general pattern is capable of extracting from a new SVO phrase. In particular, the parent phrases in Table 3 have no suffix (e.g. the *need* ends at the end-of-phrase). However, the pattern is still capable of matching correctly on a new SVO phrase where the need appears in the middle of a phrase with a non-empty suffix, "thanks to cameara" (the typo is in the original review text).

EVALUATION

We evaluate our approach on a hand-labeled set of data using three different approaches. First, we conduct precision and recall experiments as an absolute measure of performance. Second, we compare needs extracted using our automated techniques to needs generated using more traditional methods. In particular, we compare our automated identification of needs to existing consumer surveys and professional buying guides. Third, we use multi dimensional scaling to visualize the product space in terms of needs as articulated by actual users.

Data

We began with a set of 9700 digital camera reviews representing 1097 distinct product IDs collected from Epinions.com in January 2008. Using MontyLingua (Liu 2004) for linguistic preprocessing, we translated the reviews into 590,000 SVO triples. Of this set, 3,041 SVO triples were randomly selected and manually labeled by two independent coders. The Kappa of 0.4 between the coders represents "moderate" agreement. Therefore, we constructed our training/testing set from the "union" of both coders resulting in 345 distinct needs from 342 SVO triples (some triples included more than one need). Taken as a whole, the sentences were drawn from 35 distinct reviews representing 16 different products and 5 different brands. Within the labeled data, 68% of all needs appeared in SVO triples whose subject (S in the SVO triple) contains a personal pronoun (I, our, he, she, etc.) or the proper noun matching a product/brand name (canon, nikon, etc.).

Precision and Recall

For our initial evaluation, we randomly divided the labeled data into a training-set and a hold-out test-set. In accordance with our linguistic pre-processing, only SVO triples satisfying the s-filter (i.e. the subject contains a personal pronoun or a proper noun representing the product or brand) represented positive examples for training purposes. Put differently, only 235 of the 345 labeled needs were therefore eligible for training. However, all 345 labeled needs were used in measuring precision and recall. We performed 5-fold cross-validation and report recall scores for both a 60/40 and an 80/20 training/test split. Results are reported in row 1 of Table 4.

Table 4. 5-fold Cross Validation

Train/Test	60/40	80/20
s filter on training data	.58	.66
no filter on training data	.59	.63

Recall that because of our conservative decision rule based upon strictly improving the Laplacian, precision remains at 1. However, not surprisingly, recall improves with a larger training/test split. While the recall is not large relative to previous work in adaptive text extraction (Soderland 1999), the comparisons are not entirely relevant. Prior applications of adaptive text extraction have focused on semi-structured text where values are often delimited by standard boilerplate language (e.g. seminar announcements, apartment listings, etc.) or by HTML markup (Kushmerick 1997; Muslea et al. 2001). For an application that extracts strictly from natural language, we believe that these results represent a reasonable first effort.

A limitation of any supervised learning algorithm is the challenge of generating training data. By using the s-filter, we propose a heuristic for reducing the training set size; an advantage for any supervised approach. However, to test the impact of using more limited training data, we re-ran our 5-fold cross-validation without first applying the s-filter. Results are reported in Table 4 row 2. Extraction without the s-filter does not perform significantly better than extraction with the s-filter providing at least preliminary evidence that the s-filter provides comparable performance for identifying user needs at a lower training cost.

Traditional Methods

The precision and recall experiments suggest that it is possible to identify user needs within online product reviews. However, we recognize that online reviews may only reflect the needs of narrow customer market segments. In an effort to explore this question, we compared our automatically extracted needs to needs generated using traditional methods. Specifically, we identified two sets of references needs. The first set of reference needs is derived from a set of four Forrester Research North American and European technology consumption user surveys. The second set of reference needs is derived from a set of five professional online buying guides. The complete list of both are available from the authors on request. Let *i* be the intersection of the review-based needs and the reference set of needs and let *r* be the reference set of needs. Then results of the comparison are reported in Table 5 as *recall = i/r*.

Table 5. Comparing Automated Needs to Traditional Methods

	Alternative Technologies	Demographics	Psychographics	Behavioralistic
Consumer Surveys	.43	.29	.50	.33
Professional Guides		.66	.54	.80

No results are reported for the "Alternative Technologies" column within Professional Buying Guides because, not surprisingly, Digital Camera Buying Guides do not discuss the use of alternative technologies. However, the reviews capture roughly two-thirds of the demographic variables (age, gender) mentioned in professional buying guides and nearly all of the behavioralistic variables (e.g. how consumers actually use the product). That only half of the pyschographic variables (describing activities and interests of the users) are reflected in consumer reviews is perhaps not surprising. Psychographic variables reflect the diversity of different user groups. Our sample set is small to begin with, and online buying guides may be targeted at specific user populations.

The contrast with the Consumer Surveys is arguably more revealing. Although recall numbers are quite low, we might interpret this as somewhat reassuring. We began by positing that automated processing of reviews could complement traditional methods for needs elicitation. Reviews are not intended to substitute for other approaches. Because the consumer surveys are more about general consumer electronics, one might expect that reviews explicitly about digital cameras would not reflect the needs in a general survey about the use of consumer electronics. A more complete analysis would compare the professional buying guides to the consumer surveys as a benchmark to assess the significance of poor recall between the user reviews and the consumer surveys.

Multi Dimensional Scaling

Although online product reviews may not capture all of the needs in the design space, does the market segment represented by online reviewers reveal a coherent set of needs that is useful for market analysis and design? We used multi dimensional scaling (MDS) to visualize the competitive market landscape. Traditional market analysis often compares products based upon their attributes and attribute values. By automatically capturing the needs from specific product reviews, it is possible to characterize the market space in terms of user needs. The resulting map not only provides competitive information about a brand in relationship to its competitors, it also reveals market opportunities. Gaps in the market space represent focal points of unmet needs. To facilitate the mapping of our review-based needs, we manually aggregated needs into a two-level hierarchy (Ulrich and Eppinger 2008). We then applied MDS to the needs from our labeled data representing 15 different products and 5 brands. A Kruskal's Stress < 0.1 suggests that two dimensions is a valid representation and is presented as Figure 4. As a practical matter, marketers or designers might then use Principal Component Analysis to identify the key needs contributing to any one dimension.

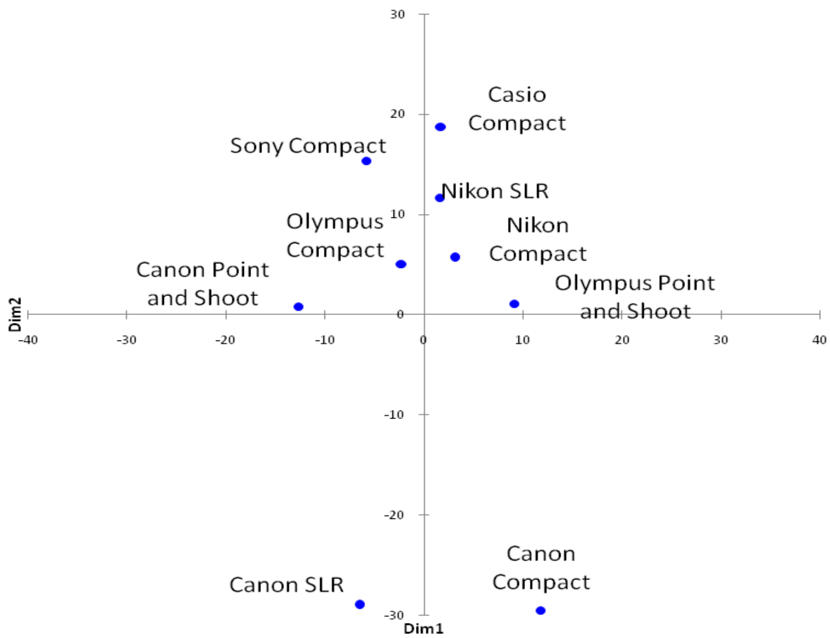


Figure 4. Multi Dimensional Scaling of the Market Based on Needs

LIMITATIONS AND FUTURE WORK

In this paper, we present an algorithm based upon adaptive text extraction and apply that algorithm to the central problem of identifying user needs. We applied the algorithm to a set of online product reviews for digital cameras and evaluated the method in three different ways. The preliminary results suggest that automated methods for analyzing product reviews hold great promise for augmenting traditional methods for assessing user needs in new product design. At the same time, as the first paper that we know of to apply automated methods to the problem of needs identification, this paper identifies a number of opportunities for future work.

First and foremost is the need for more extensive evaluation. First, the sample size consists of more than 3000 phrases. While seemingly large, it does only represent 35 distinct reviewers and product reviews. A larger sample size is necessary to establish robustness. Moreover, applying the method to other product domains and even to services is important. Prior research studying the economic impact of product reviews (Chevalier and Mayzlin 2006) suggests that different product domains may produce very different results. Although this method for identifying needs in product reviews could prove useful even if it were only applicable in narrow product domains, understanding the domain limitations is important. We are currently evaluating product reviews in several other consumer electronics domains as well as in travel services.

Beyond further testing, there are a number of possible avenues for improving the efficiency and accuracy of our method for identifying needs. To improve efficiency, we can use self-training to further reduce the amount of manual labeling required to initialize a supervised learning algorithm. Alternatively, we have begun to explore unsupervised techniques for identifying user needs through the use of language modeling and graph-based methods (Lee 2007).

To improve accuracy, we must consider parsing *within* and *between* fragments. *Within* fragments, there is the possibility of multiple LCS between two parent patterns. Heuristic search processes such as the beam search employed by Califf and Mooney (1999) are one approach. *Within* and *between* fragments, negation is a possibility (e.g. a reviewer expresses that something is not a need). However, techniques for addressing negation are already well studied within the literature on extracting product features (Hu and Liu 2004). Moreover, we believe that negation is not a concern for our specific challenge to identify an aggregate set of user needs. Even if a specific review (author) claims not to share a need, the negation declaration implies that the need could exist for some other user. Finally, because our analysis focuses on individual SVO phrases, we may miss needs that are expressed *between* or across separate SVO phrases. Compound expressions or co-reference resolution (e.g. someone writes "this is important to me" where "this" is defined elsewhere) are two such examples. First, we can extend our approach in a hierarchical fashion to learn extraction patterns that explicitly match between multiple SVO phrases (Muslea et al. 2001). As a secondary consideration, we can also empirically determine how frequently needs expressed between phrases in one review also appear within a single SVO phrase in a different review. By aggregating over all reviews, we hope to compensate for extraction errors in any one review.

Identifying needs are only one part of the overall process, however. Equally challenging is the problem of clustering those needs into coherent groups. In the marketing and product design communities, researchers have made several attempts to automatically cluster needs elicited from focus groups and interviews (Urban and Hauser 1993). However, the development of a practical method has so far proven intractable. In practice, professionals continue to cluster needs manually (Ulrich and Eppinger 2008). Clustering needs is integral to additional sensitivity analysis that we are currently undertaking. Specifically, how many reviews are enough? Because there are many different ways of articulating similar needs, clustering needs is critical to understanding the diminishing marginal returns to processing additional reviews or gathering additional needs. In their seminal work on the Voice of the Customer, Griffin and Hauser mapped the marginal returns for the traditional methods of gathering reviews (Griffin and Hauser 1993). Likewise, we are interested in identifying the minimal number of reviews required to capture a threshold percentage of the overall space of needs (recognizing the potential limitations to generalizing such results because of the possible domain specificity noted earlier).

Moreover, automated clustering would greatly facilitate two complementary lines of analysis. First, we could more easily analyze changing needs over time. For example, Kano defines a hierarchy of needs from "delights" to "competitive" to "must have's" (Urban and Hauser 1993). By tracing changes in needs over time, we could help designers identify these boundaries in addition to possibly highlighting emerging, latent needs. Second, we have begun to analyze needs for the same products but from different review sources. Users from different review sites may represent different consumer groups such as lead users (von Hippel 1986). Even if we do not capture the entire design space, following the evolution of needs within specific communities can highlight new market opportunities or highlight the transition over time between Kano's classes of needs (von Hippel 1986) .

Finally, needs discovery is only the first step in the product design process. However, automating the process of needs discovery may also lead to methods for facilitating additional steps in the design process. In the user-pull design process, Quality Function Deployment (QFD) begins with user needs and uses a matrix to map user needs to product attributes. By modeling user reviews as a knowledgebase, we can mine user reviews for association rules relating user needs to product attributes (Lee 2004). In addition to traditional user-pull, a needs-based analysis may facilitate the identification of new product development opportunities by finding opportunities to apply or integrate existing products in novel ways (Chen et al. 2004). Collaborative filtering techniques are currently used to match users based upon similar purchase histories or to match products based upon similar purchasers. However, by representing products as a vector of needs, we can match products that satisfy similar or complementary needs.

ACKNOWLEDGMENTS

This work was supported by the Fishman-Davidson Center for Service and Operations Management and the Mack Center for Technological Innovation at the Wharton School of the University.

REFERENCES

Adomavicius, G., Sankaranarayanan, R., Sen, S., and Tuzhilin, A. "Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach," *ACM TOIS* (23:1), January 2005, pp 103-145.
Archak, N., Ghose, A., and Ipeirotis, P. "Show me the Money! Deriving the Pricing Power of Product Features by Mining Consumer Reviews," in: *KDD*, 2007.
Califf, M.E., Mooney, R.J., and Cohn, D. "Relational Learning of Pattern-Match Rules for Information Extraction," in: *AAAI*, 1999.
Chen, G., Kimbrough, S.O., and Lee, T.Y. "Automated Support for Product Application Discoverry," in: *WITS*, 2004.
Chevalier, J.A., and Mayzlin, D. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), August 2006, pp 345-354.
Das, S.R., and Chen, M.Y. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science* (53:9) 2007, pp 1375-1388.
Ghose, A., Ipeirotis, P., and Sundararajan, A. "The Dimensions of Reputation in Electronic Markets," New York University, 2006, p. 32.
Griffin, A., and Hauser, J.R. "The Voice of the Customer," *Marketing Science* (12:1), Winter 1993, pp 1-27.
Hu, M., and Liu, B. "Mining and Summarizing Customer Reviews," KDD04, ACM, Seattle, WA, 2004.
Kushmerick "Wrapper induction for information extraction," in: *IJCAI*, 1997.
Lee, T. "Use-centric mining of customer reviews," in: *WITS*, Washington, D.C., 2004, pp. 146-151.
Lee, T. "Needs-Based Analysis of Online Customer Reviews," in: *ICEC*, Minneapolis, MN, 2007.

Liu, B., Hu, M., and Cheng, J. "Opinion Observer: Analyzing and Comparing Opinons on the Web," in: *WWW*, ACM, Chiba, Japan, 2005.
Liu, H. "MontyLingua: An End-to-End Natural Language Processor with Common Sense," 2004.
Muslea, I., Minton, S., and Knoblock, C.A. "Hierarchical Wrapper Induction for Semistructured Information Sources," *Autonomous Agents and Multi-Agent Systems* (4:1/2) 2001, p 21.
Nasukawa, T., and Yi, J. "Sentiment Analysis: Capturing Favorability Using Natural Language Processing," in: *K-CAP`03*, ACM, Sanibel Island, Florida, 2003, pp. 70-77.
Orman, L. "Consumer Support Systems," *Communicatons of the ACM* (50:4), April 2007, pp 49-54.
Pavlou, P.A., and Dimoka, A. "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation," *Information Systems Research* (17:4) 2006, pp 392-414.
Popescu, A.-M., and Etzioni, O. "Extracting Product Features and Opinions from Reviews," HLT-EMNLP, 2005.
Randall, T., Terwiesch, C., and Ulrich, K.T. "User Design of Customized Products," *Marketing Science* (26:2), March-April 2007, pp 268-280.
Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H., and Jin, C. "Red Opal: Product-Feature Scoring from Reviews," in: *ACM EC*, San Diego, CA, 2007.
Shankar, V., Sultan, F., and Urban, G. "Online Trust: a Stakeholder Perspective, Concepts, Implications, and Future Directions," *Strategic Information Systems* (11), 2002 2002, p 19.
Soderland, S. "Learning information extraction rules for semi-structured and free text " *Machine Learning* (34) 1999, p 39.
Turney, P. "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," ACL, Philadelphia, PA, 2002, pp. 417-424.
Ulrich, K.T., and Eppinger, S.D. *Product Design and Development*, (4 ed.) McGraw-Hill/Irwin, New York, 2008, p. 384.
Urban, G.L., and Hauser, J.R. *Design and Marketing of New Products*, (2nd ed.) Prentice Hall, 1993, p. 701.
Urban, G.L., and Hauser, J.R. ""Listening In" to Find Unmet Customer Needs and Solutions," MIT, Boston, MA, p. 36.
von Hippel, E. "Lead Users: A Source of Novel Product Concepts," *Management Science* (32:7) 1986, p 14.